ON THE SPECIFICATION OF THE TIME STRUCTURE OF ECONOMIC RELATIONSHIPS

John Bossons

01-62

01-62

May 25, 1962

ON THE SPECIFICATION OF THE TIME STRUCTURE

OF ECONOMIC RELATIONSHIPS[1]

John Bossons

School of Industrial Management
Massachusetts Institute of Technology

In order to estimate the parameters of predictive models of an economic time series it is necessary to specify the form of the stochastic process which generates the observed values of the series. This paper suggests using information on contiguous changes in the observed series in such structural specification. A variable summarizing such information is defined, and the behavior of this variable is examined for various stochastic processes which may generate an observed series. This variable is defined so as to enable information on the relationship between contiguous changes to be obtained in terms of this variable either from samples of observed realizations or from subjective data obtained from decision-makers.

1. Introduction

Most if not all economic variables depend on other variables. Even at infinite cost, however, such dependence cannot be completely specified -- the result of the economic analog of the Heisenberg uncertainty principle that it is impossible to observe all elements of a situation without the fact of observation

[1] I am indebted to a number of my colleagues at M.I.T. who have commented on various drafts of this paper. In particular, I should like to acknowledge my gratitude to Paul Cootner, Edwin Kuh, Paul Samuelson, William Steiger, and Bernt Stigum for the very helpful suggestions all of them have made. Donald Farrar of the University of Wisconsin, Michael Davis of the U.S. Navy David Taylor Model Basin, Donald Hester of Yale University, and Hendrik Houthakker and John Lintner of Harvard University have also made a number of very useful comments. This paper owes much to their suggestions. Some of the numerical analyses reported in this paper were carried out by William Steiger, to whom I am moreover indebted for pointing out an important algebraic error in an earlier draft. Computations were supported by a grant from the Sloan Research Fund of the M.I.T. School of Industrial Management, and were carried out on the School's IBM 1620 and on the IBM 7090 of the M.I.T. Computation Center.

affecting at least some of the observed elements. Consequently such dependence can only be perceived in gross, with detail masked by a mist of minor unspecified influences. Taking the cost of specification into account, the mist quickly tends to become a dense fog obscuring even major relationships among variables. "The" econometric problem is to discern such major relationships through the cloud of other influences which affect a body of data.

In this paper I shall be concerned with a subproblem within "the" econometric problem: namely, with the specification of the time-dimensional structure of relationships among variables. In the past, it has commonly been assumed that this structure can be adequately described by relatively simple lag structures between variables or first differences of variables. I believe that in many circumstances use of such simple models of time-dimensional structure can lead to substantial mis-specification of relationships, and therefore propose to examine the time-structural aspect of the general specification problem more closely.

The general specification problem really involves two specification problems: (1) specifying a conceptual model of the relationships between a variable and other variables which influence it, and (2) specifying a model of the correspondences between the variables specified in the conceptual model and variates observable in the real world. Both problems are involved in the subproblem of specifying time-dimensional structure, which can be defined more specifically as the problem of specifying the time-pattern of changes in a variable which are generated by a given change in another variable. It is necessary both to specify a conceptual model describing the time-pattern of this dependence and to specify the empirical meaning of the changes in terms of

which the structural model is defined. Since changes generated by an identifiable change in another variable cannot be observed apart from changes generated by other changes in that variable or by changes in other variables, the second part of the general specification problem is as critical as the first. In order to identify the time-dimensional structure of a dependence of one variable on another, it is necessary first to specify the time pattern of the relationship between the dependent variable and independent variables not included in the model to be specified.

In this paper, I shall be concerned exclusively with this "prerequisite" specification, which may be termed the specification of the time-dimensional structure of unspecified relationships between a variable and other variables. I shall discuss the nature of this specification in more detail in the next section, and will suggest a new method of analysis which can be used both to supplement current analyses of observed values of a time series and to incorporate subjective data. I shall then go on in further sections to develop this method of analysis.

The pragmatic justification for developing such analytic techniques need hardly be discussed. The importance of specifying the time-dimensional structure of economic relationships is becoming increasingly apparent for numerous economic models which attempt to describe the effects of actions taken by decision units in the economy. It is not difficult to enumerate a by no means exhaustive list of examples of problems in which such specification is important: explaining the behavior of aggregate consumption, analyzing the nature of reactions to changes in the price of a particular commodity, describing the dividend policies of corporation managers, deciding upon the storage capacity of a dam to be built for flood-control purposes, forecasting changes in a manufacturer's sales or inventory, and so forth. In all these applications, it is essential to

specify, insofar as is possible, the time-dimensional character of the variables which enter each model as inputs to the decision process, quite apart from the problem of specifying the nature of the relationships between those variables and the decision variable.

2. Analyzing the time structure of "independent" variables

In order to deal with the simplest case, I shall define the "prerequisite" time-structure specification problem in terms of a variable which is not clearly dependent on any other variables, so that changes in that variable are determined by influences which are unidentified and hence unpredictable. In the face of the unpredictability of these influences, the "safest" specification that can be made -- safest in terms of minimax criteria -- is to postulate that changes in the sum of their effects are realizations of a pure random process. There may well be error in such specification, of course. But even were it possible to know with certainty the magnitude of such error in any given period, the most accurate specification that can be made in the absence of information about the direction of that error still would be merely to flip a fair coin to "predict" in which direction the change of known magnitude will be.[1]

If some of the influences upon a variable __are__ identifiable, then clearly it would be erroneous to postulate that changes in the sum of the effects of all influences were random, since it would then be possible to [illegible] changes [illegible] influences. It may be possible, for instance, to [illegible] orders for a company's product are partially determined by [illegible]

---

[1] This reasoning implies that the Savage axioms are valid for normative purposes. On this point, see the discussion of the Knightian distinction between "measurable and 'unmeasurable' (sic?) uncertainty in Ellsberg [21], along with Howard Raiffa's perceptive comments in [56].

or in the company's expenditures on advertising. If so, it should then prove possible to improve predictions of changes in the new order rate over those obtained by simply assuming all future changes in influences on orders to be realizations of a random process and on that basis extrapolating the stochastic process of which new orders are specified to be realizations. To do so, it will be necessary to specify the time-dimensional structure of the relationship between the variable and the identified influence as well as of the relationship between the variable and unidentified influences. I shall not discuss such simultaneous specification and estimation in this paper, but will instead restrict the analysis to specification of the time-dimensional structure of the relationship between a variable and influences which are assumed to be all unidentified.[1] I shall thus restrict myself to the specification aspects of the classic problem of time series analysis.

Since any time series is but a sample realization of some stochastic process, the essence of the problem of either predicting or describing such a series is to describe the underlying stochastic process. Since any stochastic process can in turn be defined as a transformation of a purely random process, time series analysis may thus be defined as, in essence, the estimation of the nature of this transformation from a set of observations on the economic time

---

[1] Where the time pattern of effects of an identified influence cannot be specified apart from that of unidentified influences, information about forthcoming events can still be incorporated into a forecast by simply equating the expected value of a future "random" shock to the expected total effect of the identified influence then current and assuming the time pattern of effects of that influence to be the same as that of unidentified influences. By specifying a model of a time series' structure, it is thus possible for a manager to incorporate information about future shocks explicitly into that model's forecasts, rather than using that information as the basis either for "fudging" the model's prediction or for avoiding the model entirely.

series.[1] Definition of this transformation provides a formal definition of the structure of the relationship between a variable and the sum of all influences upon it where those influences are unidentified.

The foundation of classic time-series analysis may thus be viewed as the postulated existence of a purely random process defining a sequence of random shocks $\{\zeta(j);\ j \in T\}$ which are independently, identically, and normally distributed for all values of the index set T, with zero expected value and unit variance. This set of random shocks by definition is a strictly stationary stochastic process, corresponding to what (in the continuous case) a communications engineer would term "pure white noise." Defining, then a transformation Ψ (not necessarily linear) such that

$$(2.1) \qquad \{X(t);\ t \in T\} = \Psi \{\zeta(j);\ j \in T,\ j \leq t\}$$

it is necessary to be able to estimate Ψ from the observed realization of the stochastic process X(t). If this transformation is linear and X(t) is stationary, then spectral analysis provides a relatively powerful means of estimation.[2] (Satisfactory estimation techniques have not yet been developed to estimate non-linear transformations.)

---

[1] Time series analysis may be <u>defined</u> in this manner. Any stochastic process <u>can</u> be written as a transformation of a purely random process. To state that this definition is useful is another matter, however, involving as it must some statement about the inverse realizability of the process as so defined. It can be conjectured that it is useful to regard the prediction problem as that of estimating the nature of the transformation defined in equation (2.1). Norbert Wiener has taken essentially this approach ir [70]. As Davis [17][18] has demonstrated, it can be shown that optimal linear prediction can be regarded, as conjectured, as a measurement of state variables in a white noise excited model and of their initial condition decay. Extension of this result to non-linear systems is not facile. However, since all models dealt with in this paper are linear in the purely random process, Davis' result is valid for these models.

[2] See for instance Grenander and Rosenblatt [29], Parzen [54, chapter 5], Rosenblatt [60, chapter 7], and Whittle [69]. Stigum [65] generalizes this means of estimation to include certain classes of non-stationary processes. For a brief exposition of the interrelationship between the statistical analysis of (stationary) time series and the theory of stochastic processes possessing finite second moments, see Parzen [55] or Bartlett [5].

However, as Bernt Stigum has shown in an important unpublished paper [65], it is not generally possible to specify a unique transformation from the spectrum derived from a given set of time series observations. Bayesian estimation is consequently necessary in the analysis of time series; that is, even given that such a transformation would be linear, it is necessary to specify the form of the relevant transformation on the basis of economic theory and other "prior" knowledge.[1] The immediate purpose of this paper is twofold: (1) to classify the stochastic processes implicitly specified by various economic models in terms of narrower specifications of the Ψ -transformation defined in (2.1), and (2) to suggest a means of utilizing not only objective sample data in specifying Ψ but also expectational and interview data on decision-makers' own evaluations of the nature of the stochastic processes underlying time series in which they are interested. I shall not concern myself with the problem of estimating Ψ once the nature of Ψ is known, but shall confine myself purely to the specification problem. Though both goals stated above are motivated by the need to develop better means of specifying the time structure of economic variables, the second should as a by-product also provide the basis for analyzing hypotheses about decision-makers' anticipations.

A substantial body of information on economic time series is available only in terms of the relationship between successive changes in an observed time series. Much expectational information, for instance, is perceived in this form.

---

[1] This position is of course in conflict with the methodological structure of Milton Friedman [27], as indeed is any Bayesian position. Perhaps the most useful function of the pathbreaking papers of Yule [74] and Slutzky [63] on the causation of so-called "cycles" is to warn of the great danger involved in relying upon models of economic time series which are not based on previously-derived "realistic" models of such series' structure.

While such data is not sufficiently precise to be worth subjecting to very heavy analysis, it may contain information which can shed light on the structure of the time series. It should therefore be useful to develop for both subject and sample data a means of evaluating the correspondence between relationships between successive changes and the time-dimensional structure of a variable.

I suggest that perhaps the most fruitful way of examining this correspondence is to define a variable which will summarize the relationship between successive changes, and then to analyze, for different time series' structures, the relationship between this variable and different lengths of the intervals over which changes are defined. For convenience, I propose to think in terms of changes centered around some fixed point t. Let Δ_R denote change in X over the interval between t-R and t for an arbitrarily-chosen R, and let Δ_F denote change in X between t and t+F for an arbitrary F. (The intervals between t-R and t and between t and t+F may be termed the record and forecast horizons, respectively). Define a variable γ_{FR} such that a decision-maker's conditional expectation for Δ_F given Δ_R can be written as

(2.2) $$E\,[\Delta_F \mid \Delta_R] = \gamma_{FR}\,(\Delta_R).$$

I propose to use this variable to describe the relationship between the successive changes denoted by Δ_F and Δ_R. Other models of this relationship might be specified. My ground for choosing this model is the purely pragmatic justification that virturally all quantitative data can be expressed in terms of this model and that much qualitative data can be expressed only in terms of it. By establishing the properties of γ_{FR} for different specifications of the Ψ-transformation, it will thus be possible to use qualitative information such as decision-makers' subjective estimates of γ_{FR} for different F and R which are formed on the basis of their experience and their familiarity with the nature of the influences upon the variable

being examined. It will at the same time be possible to estimate the relationship between γ_{FR} and the values of F and R directly from historical records of observations of a time series. Such direct estimation is used in applications discussed in Section 6 below. But such direct estimation will be of limited value where the relevant historical record is small.[1] It should prove useful in many applications to be able to incorporate subjective information obtained in a form consistent with that of information derived from sample data.[2]

Both $E[\Delta_F \mid \Delta_R]$ and Δ_R may be defined quite broadly. The values which they may take on may be the set of all real numbers; they may comprise only a selected set of integers; or they may simply be the trichotomy consisting of "up," "down," and "no change." In this paper I shall assume both $E[\Delta_F \mid \Delta_R]$ and Δ_R to be real numbers; γ_{FR} is thus real-valued. If by contrast observations of $E[\Delta_F \mid \Delta_R]$ and Δ_R are limited to statements of direction of change rather than

---

1 The historical record may be small either because past data is available only in very limited quantity or because portions of the historical data are known to be influenced by special factors which would introduce estimation bias if such data were incorporated into the sample from which γ_{FR} is estimated. The notion of a "relevant"historical record is a vague concept: what is meant is a historical record over a period in which the parameters of a variable's time structure can be assumed (1) to be stable, and (2) to accord with those of the variable in subsequent periods for which structural estimates are to be used for forecasting purposes. A "relevant"record so defined should exclude observations known to contain misleading information; on this, see Frank Fisher's pungent remarks in [23, chapter 1].

2 Since both subjective and historical information are then summarized in the same set of parameters (γ_{FR}), an estimate taking both sets of information into account may be formed by simply combining the two sets of parameters estimates in a weighted average. Using S, H, and P as superscripts respectively to denote subjective, historical, and pooled estimates of γ_{FR}, an analyst could set

$$\gamma_{FR}^{P} = w\,\gamma_{FR}^{S} + (1-w)\,\gamma_{FR}^{H},$$

where the weights w and (1-w) reflect the analyst's evaluation of the relative information content of γ_{FR}^{S} and γ_{FR}^{H}. (This procedure is analagous to the combination of prior and posterior probabilities discussed in e.g. [57, chapter 12]and [61]; it is not the same thing, however, since the subjective estimates being combined with sample estimates are those of individuals other than the analyst and are themselves derived by interview or other sample survey techniques.)

real numbers, as in certain anticipations surveys such as those conducted by Dun and Bradstreet or the IFO-Institut für Wirtschaftsforschung, the analysis developed in this paper may be applied to such observations by equating each element of the trichotomy of directions with three intervals defined on the range of real numbers.[1]

In order to analyze the behavior of γ_{FR} for different stochastic processes, I shall define it more narrowly as the expected value (measured over an infinite ensemble of realizations of the process) of the slope of the regression of Δ_F on Δ_R constrained to pass through (0,0), namely, as the expected value of

$$\hat{\gamma}_{FR} = \frac{\sum \Delta_F \Delta_R}{\sum \Delta_R^2} \tag{2.3}$$

This definition is easily applicable to observed data.

Empirically derived data on γ_{FR} will of course be subject to error. Such errors can be explicitly introduced for estimates of γ_{FR} obtained from objective data by stating equation (2.2) in the following form:

$$\Delta_F = \gamma_{FR} (\Delta_R) + u_{FR} \tag{2.4}$$

where u_{FR}, the difference between Δ_F and $E [\Delta_F \mid \Delta_R]$, includes the difference between the true values of Δ_F and the conditional expectation as well as errors of observation in each. If u_{FR} satisfies the standard Markov theorem assumptions of

---

[1] The two numbers which demarcate the three intervals (or subsets of the set of all real numbers) corresponding to "up," "down," and "no change" may be thought of as perception thresholds. An illuminating discussion of such thresholds in terms of response functions is contained in Theil [67, pp. 196-199]. While some specification error is introduced by considering thresholds to be fixed points rather than continuous functions à la Theil, such error should in most cases be of minor importance.

independence, zero expected value, and homoscedasticity, then $\hat{\gamma}_{FR}$ is an unbiassed and by minimum-variance criteria a "best" estimate of γ_{FR}. It should be noted that the relationship between γ_{FR} and values of F and R, which will be denoted by the function γ (F.R), is intimately related to the form of the autocorrelation function R(L) defined over lags of L = 0,1,2,.... Specifically, if u_{FR} satisfies the Markov theorem assumptions then

$E\Delta_F\Delta_R = \text{cov}(\Delta_F\Delta_R)$, $E\Delta_R^{\;2} = \text{var}(\Delta_R)$, and

$$(2.5) \qquad \gamma_{FR} = E\,\hat{\gamma}_{FR} = \frac{[R_F + R_R] - [R_0 + R_{F+R}]}{2[R_0 - R_R]}$$

Consequently, where it can be assumed that u_{FR} satisfies the least-squares assumptions, I will, in deriving the properties of $\gamma(F,R)$ for various stochastic processes, really be indirectly deriving the nature of the autocorrelation function for those processes.[1] In such cases, if the autocorrelation function for a process has previously been derived, I shall derive the properties of $\gamma(F,R)$ in terms of that function.

I have justified estimation of γ_{FR} primarily in terms of the usefulness of developing analytic techniques that can be applied to subjective as well as numerical data. However, it should not be assumed, where only historical data are available, that estimation of $\gamma(F,R)$ is redundant once the autocorrelation function has been estimated in cases where u_{FR} satisfies the Markov assumptions. Though both R(L) and $\gamma(F,R)$ summarize the same information in such cases -- both

---

[1] It should be noted that the serial coefficient estimates of the R_L implicitly specified by (2.5) as defining the estimates of γ_{FR} for a given N-observation sample are not identical to the estimates of the same R_L which will typically be obtained from that sample since the estimates specified in (2.5) are defined over a subsample including but N-F-R observations compared with the N-L observations on which R_L is typically estimated.

functions are estimated so as to filter out noise from the set of serial coefficient estimates of R_L obtained from a given sample -- each summarizes that information in a different manner and so contains different information about the information being summarized. Because of their interdependence, simultaneous estimation of the two functions should consequently lead to better estimation of both.

I shall not deal with possible sources of bias in estimates of γ_{FR} obtained from either subjective or objective sample data, even though, in the absence of proof of the non-existence of such error, estimates of the potential likelihood of such error must be made in order for $\gamma(F,R)$ to be a useful diagnostic tool.[1] Where γ_{FR} is derived from observation of decision-makers' qualitative appraisals of $E[\Delta_F \mid \Delta_R]$, for instance, the associated errors denoted by u_{FR} clearly may not satisfy the least-squares assumptions, since the decision-makers' conditional expectations may very well not be unbiased -- or in any way be "best estimates."[2] Nevertheless, I shall not deal with this problem in this paper, confining myself to an evaluation of the nature of the function $\gamma(F,R)$ which is implied by various specifications of the Ψ-transformation, defining this function as the relationship between F,R, and the expected value of the estimate $\hat{\gamma}_{FR}$ defined in equation (2.3).

---

1 Sampling properties of estimates obtained from objective sample data will be further analyzed in [11] both for the constrained regression defined in equation (2.2) above and for unconstrained linear regressions of Δ_F on Δ_R.

2 It is possible, of course, to define unbiased forecasts as "rational" and to assume that businessmen, consumers, and other decision-makers in the economy are rational in the sense of this definition. This is not to say that this procedure is reasonable, however. Cf. [46], [48], and the comments on this approach in [10, esp. pp. 7-8] and [8].

Nor shall I be concerned with analyzing $\gamma(F,R)$ in terms of the path traced through time by $X(t)$, even though this problem has historically been of interest in the analysis of business fluctuations. Since even a random walk can generate time paths that are sufficiently "cyclical" to approximate the fluctuations present in most economic time series,[1] the construction of models to propagate such "cycles" is not a particulary interesting activity.

Having indicated what I will not do, let me delineate what I will do. I shall proceed to analyze the relationship between γ_{FR}, defined as the expected value of the estimate $\hat{\gamma}_{FR}$ defined in (2.3), and the horizon parameters F and R for different discrete-process models of $X(t)$. In Section 3, I shall do this for a simple stochastic process which has recently seen fairly widepsread use in economic models. In Section 4, I shall then use the model defined in Section 3 as the second stage of a two-stage model of $X(t)$, defining the first stage of this model as an intermediate stochastic process which is not necessarily a pure random process but which can be written as a relatively straightforward transformation of one which is. (Where the intermediate process <u>is</u> purely random with zero mean and constant variance, the model of Section 4 reduces to that of Section 3). In Section 5, I shall turn to alternate ways of specifying the form of a second stage model which is linear in the elements of the intermediate stochastic process, and shall analyze the behavior of γ_{FR} in these models.

---

[1] See for instance Feller [22, chapter 3] and Kerrich [36]. An earlier example is provided by Working's 1934 paper [73]. The conception of the effectiveness of random shocks in generating "cycles" is due to Eugen Slutzky [63], G.U. Yule [74], and Ragnar Frisch [28]. The papers of Slutzky and Yule, both published in 1927, showed that a summation of random shocks produces a series which can be well described <u>ex post facto</u> by a Fourier equation (i.e., as a summation of cycles of varying frequency) even though this equation does not predict change in the observed time series.

An illustrative application of the analyses developed in Sections 3 through 5 will be presented in Section 6.

The models developed in Sections 3 through 5 are all linear in the underlying pure random process, and are hence special cases of the Yule-Slutzky processes described e.g. by Wold [72]. I suspect, however, that they are adequate to describe most economic time series. As will be seen, the stochastic processes analyzed in Sections 3 through 5 are not all stationary - some, indeed, not even in thw weak sense of covariance stationarity. While this does not make estimation easier, it does at least suggest that the range of processes described below may be sufficient to include most time series of interest to economists and managerial scientists.[1]

---

[1] Like the optimality of Ψ-transform estimation postulated supra, this claim of broad applicability is a conjecture, with little empirical evidence presently available to validate it. Mandelbrot [43] has suggested that many economic time series are realizations of a quite different class of processes (Pareto-Levy processes) which, in contrast to the processes specified below in Sections 3 through 5, do not satisfy the conditions for applicability of the Central Limit Theorem since finite second moments do not exist for such processes. I have not attempted to analyze $\gamma(F.R)$ for such processes. So far, little empirical evidence has been presented to indicate whether any economic time series is realized from a Pareto-Levy process, or, for that matter, any other process. Cf., however, Mandelbrot [42][44], who has presented some evidence indicating that both income distributions and the movement of commodity prices may be determined by Pareto-Levy processes.

Another class of processes which I have ignored in Sections 3 through 5 is the set of processes which are defined as the sum of two orthogonal components: a deterministic mean-value function $M(t)$, defining a systematic oscillation and/or polynomial trend over $t \in T$, and a stochastic process $X(t)$. Specification of such processes forms the basis for the decomposition approach which characterizes much economic time series analysis. See for instance the orthogonal polynomial approach of Hald [30] and the alternative attempts to measure $M(t)$ by multiple smoothing suggested by among others Maverick [45] and Brown and Meyer [13]. Wold [72, Theorem 7] has shown that any stationary process can be decomposed into $M(t)$ and $X(t)$ and that moreover $X(t)$ is then linear in the pure random process $\zeta(t)$. Analyses such as [30], [45], and [13] usually assume that $X(t) = \zeta(t)$; for a general analysis of Wold-decomposable processes are Rosenblatt [59]. I suspect, however, that most if not all economic processes are characterized by mean-value functions identically equal to zero. If this is so, an attempt to measure $M(t)$ rather than the structure of $X(t)$ is not only a misorientation of focus but also is likely to result in the type of misspecification against which the Yule and Slutzky papers warn. The application to common stock prices in [13] of the multiple-smoothing formulae advanced by Brown and Meyer may be a good example of the dangers of such an approach if common stock prices are really a random walk as suggested by [1], [35], and [58].

3. A simple stochastic process

The stochastic process implicitly specified by a large number of models advanced in recent economic literature can be derived from the following special assumptions: (1) each realization of the process is determined up to a realization of a pure random process by the effects of previous realizations of that pure random process, and (2) the effect of any realization of the pure random process is identical in all periods subsequent to that realization and furthermore is proportional to the size of the realization. This derivation is demonstrated in section 3.1. The nature of the relationship between Δ_F and Δ_R in time series generated by such a stochastic process is then examined in section 3.2. As will be shown in section 3.3, the stochastic process defined in section 3.1 can be regarded as a linear combination of two other, more specialized stochastic processes -- a random walk and a time series of independently and identically distributed variables -- and these special cases will be briefly discussed in sections 3.3. Some implications of the process concerning the variance of X(t) are analyzed in section 3.4.

3.1 A "proportional effect" model

It has become popular, following the example delineated by Milton Friedman, to regard a random shock as having two components, one transient and the other persistent, and to assume that the proportionate importance of each component is the same for all random shocks.[1] It is also assumed that the random shocks are themselves

---

[1] Friedman's most well-known model is that presented in [26], though essentially the same model was described in an earlier work in which Friedman and Simon Kuznets were co-authors [27]. The terminology I have adopted is different from that used by Friedman, who denoted the two components into which random shocks could be partitioned as "permanent" and "transitory." "Persistent" probably somewhat more fully connotes the necessary existence of a time horizon over which that component is defined.

realizations of a pure random process $\{\xi(t);\ t \varepsilon T\}$ with mean zero and variance b^2 -- or, in other words, that they are realizations of a process which is a linear homogeneous transformation of the process $\{\zeta(t);\ t \varepsilon T\}$ such that

(3.1.1) $$\xi(t) = b[\zeta(t)].$$

From the assumption of equal proportionate importance of the persistent component for each shock, each shock thus has an identical additive effect on the value of $X(t)$ for all $t > j$. This effect can be denoted as $\lambda[\xi(j)]$, $0 \leq \lambda \leq 1$, where λ represents the proportionate size of the persistent component of each random shock. This model implicitly assumes that the observed time series, $X(t)$, is the sum of the persistent components of all previous random shocks together with both the transient and the persistent components of the random shock occurring at t. As a consequence

(3.1.2) $$X(t) = \xi(t) + \lambda \sum_{j<t} \xi(j)$$

By substituting from (3.1.1), $X(t)$ can be defined in terms of the unit-variance Gaussian "white noise" process:

(3.1.3) $$X(t) = b\,\zeta(t) + b\,\lambda \sum_{j<t} \zeta(j)$$

The expected value of $X(t)$ given the values of the previous random shocks -- the previously determined component of $X(t)$, in other words -- is thus

(3.1.4) $$E[X(t) \setminus \zeta(j),\ j<t] = b\lambda \sum_{j<t} \zeta(j).$$

For convenience, I shall introduce some new notation. The conditional expectation $E[X(t) \mid \zeta(j),\ j < t]$ will be denoted by the symbol $\bar{\bar{X}}_t$; in addition, the time index will be subscripted so that X(t) will be denoted by X_t. Using this notation, it is evident from equation (3.1.4) that

$$\bar{\bar{X}}_j = \bar{\bar{X}}_{t-1} + b\,\lambda\,\zeta_{t-1} \tag{3.1.5}$$

so that

$$\bar{\bar{X}}_t - \bar{\bar{X}}_{t-1} = b\,\lambda\,\zeta_{t-1} \tag{3.1.6}$$

Period-to-period changes in the previously determined component of a time series are thus defined by this model to be proportional to the intervening random shock.

This model is equivalent in all respects to the exponentially-weighted moving averages which have been specified in numerous models in recent economic literature as the relevant form of a distributed lag structure.[1] This equivalence can be quickly demonstrated by showing from (3.1.2) and the definition of $\bar{\bar{X}}_t$ that $X_t = \bar{\bar{X}}_t + b\,\zeta_t$ and that equation (3.1.6) can therefore be written

$$\bar{\bar{X}}_t - \bar{\bar{X}}_{t-1} = \lambda\,[X_{t-1} - \bar{\bar{X}}_{t-1}] \tag{3.1.7}$$

Changes in the conditional expected value of X_t are thus defined as proportionate to the error associated with the immediately previous conditional expectation. Equation (3.1.7) is quickly recognizable as the definition of the "adaptive" expectations

---

[1] A partial list of recent applications of exponentially-weighted moving averages includes the analyses of investment behavior by Koyck [37], Eisner [20], and Kuh [38]; the demand analyses of Nerlove [50][51]; the analyses of aggregate consumption by Stone and Rowe [66] and Friedman [26]; and the analysis of monetary hyperinflation by Cagan [14]. A slightly different formulation of this model was postulated by Lintner [38] in explaining the relationship between dividends and corporate profits. Exponentially-weighted forecasts have been advocated as both positive and normative models of suppliers' price or sales anticipations in Arrow and Nerlove [2], Brown [12], Forrester [24, appendix E], Holt [33][34], Magee [41], Nerlove [52][53], and Winters [71].

which have been postulated by Nerlove and Cagan. Rearranging terms in this equation,

$$(3.1.8) \qquad \bar{\bar{X}}_t = \lambda X_{t-1} + (1-\lambda) \bar{\bar{X}}_{t-1} .$$

This equation is just one more way of stating that the adaptive expectations model defines "rational" conditional expectations as a linear combination of the preceding value of a precast variable and the conditional expectation of that value. Iterating within the right-hand side of (3.1.8), it is apparent that

$$(3.1.9) \qquad \bar{\bar{X}}_t = \lambda[X_{t-1} + \sum_{j<t-1} (1-\lambda)^{t-j-1} X_j].$$

The conditional expectation denoted by $\bar{\bar{X}}_t$ is thus an exponentially weighted average of all previous values of X.

3.2 Analysis of γ_{FR} for the "proportional effect" model

The relationships between successive changes in X for changes defined over different intervals is relatively simple to analyze for the "proportional effect" model in terms of the estimate of γ_{FR} defined in (2.4). First of all, it should be evident from the definition of the "proportional-effect" model that, since the effect of a random shock on subsequent values of a time series if assumed to be independent of elapsed time, γ_{FR} is unrelated to the length of the forecast horizon.[1] This can be seen by generalizing equation (3.1.5) to define the following relationship between $\bar{\bar{X}}_{t+F}$ and $\bar{\bar{X}}_t$:

[1] The fact that predictions derived from an exponentially-weighted moving average are unrelated to the forecast horizon over which the predictions are made has previously been stated by Muth [47].

$$\bar{\bar{X}}_{t+F} = \bar{\bar{X}}_t + b\,\lambda \sum_{j=t}^{t+F-1} \zeta_j \,. \tag{3.2.1}$$

Since $\bar{\bar{X}}_j = X_j - b\,\zeta$, then, denoting $(X_{t+f} - X_t)$ by Δ_F, equation (3.2.1) can be rewritten as

$$\Delta_F = b\,\zeta_{t+F} + b\,\Delta \sum_{j=t+1}^{t+F-1} \zeta_j - b(1-\lambda)\,\zeta_t . \tag{3.2.2}$$

Since Δ_R, the change in X over some interval of R periods prior to t, contains no information about ζ_j for $j > t$, it is evident that E $[\Delta_F \mid \Delta_R]$ is unaffected by ζ_j for $j > t$. Moreover, since $\partial \Delta_F / \partial \zeta_t$ is a constant, the independence of the conditional expectation from all ζ_j, $j > t$, then implies that γ_{FR} is a constant for any given value of R.

A corollary of this statement is that γ_{FR} is a function of λ. Moreover, since Δ_R is itself dependent on ζ_j for $t-R \leq j \leq t$, it is evident from the same line of argument that γ_{FR} is also a function of R. By extension of (3.1.5),

$$\Delta_R = b\,\zeta_t + b\,\lambda \sum_{j=t-R+1}^{t-1} \zeta_j - b\,(1-\lambda)\,\zeta_{t-R} \tag{3.2.3}$$

Both (3.2.2) and (3.2.3) can then be substituted into (2.4) to yield the following equation defining γ_{FR}:

$$\begin{aligned} & b\,\zeta_{t+F} + b\,\lambda \sum_{j=t+1}^{t+F-1} \zeta_j - b(1-\lambda)\,\zeta_t \\ & = \gamma_{FR}\left[b\,\zeta_t + b\,\lambda \sum_{j=t-R+1}^{t-1} \zeta_j - b(1-\lambda)\,\zeta_{t-R}\right] + u_{FR} . \end{aligned} \tag{3.2.4}$$

I shall specify that u_{FR} satisfies the usual least-squares assumptions in order to obtain the expected value of the estimate of γ_{FR} defined in (2.3). Moreover, since $\zeta(t)$ is a pure random process, E $[\zeta_t\,\zeta_j] = 0$ for $j \neq t$. Consequently, multiplying both sides of (3.2.4) by the term in square brackets of that equation and

forming expected values of the resultant product moments, we obtain

$$\gamma_{FR} = \frac{-b^2(1-\lambda)\ \mathrm{var}\ \zeta_t}{b^2\ \mathrm{var}\ \zeta_t + b^2\lambda^2 \sum_{j=t-R}^{t-1} \mathrm{var}\ \zeta_j + b^2(1-2\lambda)\ \mathrm{var}\ \zeta_{t-R}} \tag{3.2.5}$$

Since the denominator of (3.2.5) is by definition positive, then since $0 \leq \lambda \leq 1$ it is evident that γ_{FR} is always non-positive and indeed is strictly negative for $\lambda < 1$. Any time-series satisfying the assumptions of the model defined in section 3.1 is thus regressive, in the trend-reversing sense defined by analysts of anticipations surveys,[1] as an automatic consequence of the presence of a transient component in the random shocks to which the series is subject.

Since var $\zeta = 1$, (3.2.5) reduces to

$$\gamma_{FR} = \frac{\lambda-1}{2(1-\lambda)+\lambda^2 R} \tag{3.2.6}$$

so that γ_{FR} depends only on λ and R and not on b. Chart 1 portrays the relationship between $-\gamma_{FR}$ and different values of λ and R. It is evident from this equation that γ_{FR} decreases monotonically from zero to -1/2 as the transcience of the random shocks increases: the smaller the value of λ, the more evanescent the random shocks and the greater the degree of regressiveness in the resulting time series.

3.3 Two special cases

The upper bound of λ defines a special case of the model defined in section 3.1 which is of particular interest. When $\lambda = 1$, then, substituting this value of λ in equation (3.1.8),

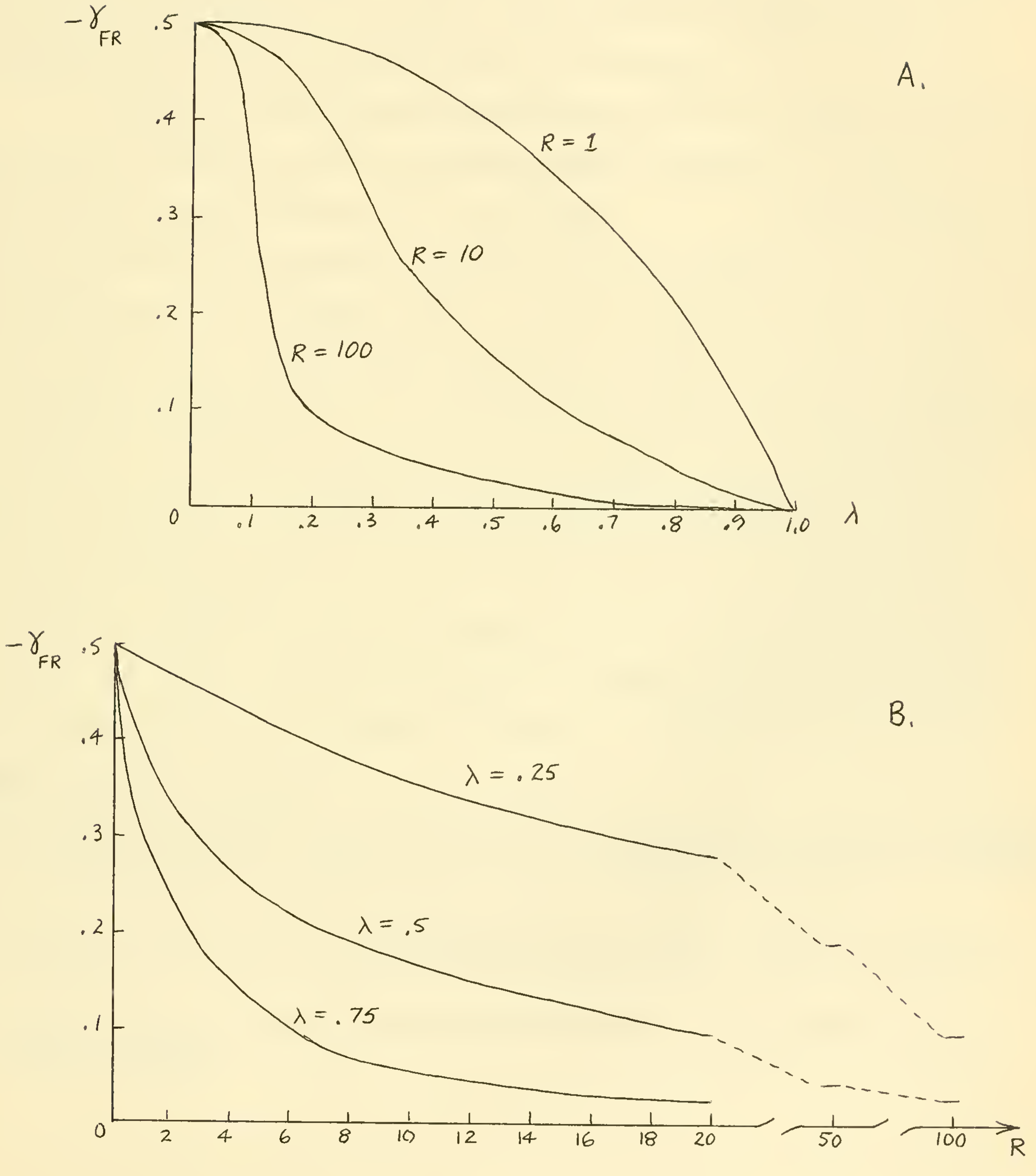
$$\bar{\bar{X}}_t = X_{t-1}. \tag{3.3.1}$$

---

[1] See for instance [10].

CHART 1

RELATION BETWEEN $-\gamma$, λ, AND R



Since $\bar{\bar{X}}_t = X_t + \zeta_t$, $E[X_t \mid X_{t-1}, \ldots] = X_{t-1}$. The stochastic process defined by this special case is thus a random walk. As might be expected, substitution of $\lambda = 1$ in equation (3.2.6) yields $\gamma_{FR} = 0$ so that, for any Δ_R, $E(\Delta_F) = 0$. In other words, knowledge of past changes in X provides no information about future changes -- a statement which is the classic definition of a random walk.

Because in a perfect market prices will be set so as to reflect the expectations about future prices held by marginal buyers and sellers, any changes in prices (apart from growth at a rate equal to the marginal time-plus-risk discount on future income) will reflect unanticipated events that, a priori, can only be viewed as realizations of a pure random process. Since the time series formed by the movement through time of a price which is determined in a perfect market will thus be a random walk if expectations are "rational,"[1] this special case of the model defined in section 3.1 is of particular interest to economists. Several authors have suggested, for instance, that even though some imperfections undoubtedly exist in security and commodity markets, prices in such markets are essentially random walks.[2] In addition, a random walk may also describe other series not determined in perfect markets. One example, constructed in terms of a highly imperfect market, might be the cumulative sales made by individual salesmen in a market where personal contacts are highly important.[3]

[1] This condition is sufficient but not necessary. Expectations need not be "rational" in the Muthian sense of being unbiassed; it is necessary merely that any systematic relationship between expectations and outcome be invariant over time.

[2] See for instance Alexander [1], Bachelier [3], and Kendall [35], and the further exposition of this view in Cootner [15], Houthakker [34a], and Roberts [58].

[3] Random shocks would in such a model represent primarily the salesman's gains or losses of contacts.

The stochastic process defined by (3.3.1) is of course an unconstrained random walk. In a number of applications, a random walk may be bounded by absorbing or reflecting barriers. A random walk with absorbing barriers (as in Russian roulette) is particularly interesting in economic applications because of the problem of evaluating the "value" of ruin.[1] While such constrained random walks are not fully described by (2.3.1), γ_{FR} is nevertheless not affected within known barriers which are sufficiently far away from most values of X_t.

A second special case of some interest is defined by the lower bound of λ. When $\lambda = 0$, then substituting this value of λ in (3.1.8) obtains

$$\bar{\bar{X}}_t = \bar{\bar{X}}_{t-1} , \tag{3.3.2}$$

so that by extension $\bar{\bar{X}}_t = \bar{\bar{X}}_j$ for all values of j. Successive values of the time series defined by (3.3.2) can thus be viewed as sample observations drawn from an identical population. That the values of these observations are independent of their ordering in time is merely another way of saying that each random shock is completely transient.

Substituting $\lambda = 0$ into (3.2.6) yields $\gamma_{FR} = -1/2$ for any value of R. While this result may not be intuitively obvious, it can be explained by considering that random variation in Δ_R may be attributed to either X_t or X_{t-R} with equal likelihood. In other words, since $\Delta_R = \zeta_t - \zeta_{t-R}$ because $\bar{\bar{X}}_t = \bar{\bar{X}}_{t-R}$ (both ζ_t and ζ_{t-R} being identically distributed with zero expected value), $E[\Delta_F | \Delta_R]$ may be regarded as being, with equal likelihood, any linear combination of

---

[1] See for instance the model of dividend policy cast in these terms contained in Shubik [62]. The classical "gambler's ruin" problem is discussed in Feller [22, pp. 313-318].

$E[\Delta_F \mid \Delta_R = \zeta_t]$ and $E[\Delta_F \mid \Delta_R = -\zeta_{t-R}]$. Since, from (3.2.2),

$$E[\Delta_F \mid \zeta_t] = -\zeta_t \tag{3.3.3}$$

and

$$E[\Delta_F \mid \zeta_{t-R}] = 0, \tag{3.3.4}$$

it is thus evident that $E[\Delta_F|\Delta_R] = -1/2\ \Delta_R$ so that $\gamma_{FR} = -1/2$. It is in any case clearly evident that $\gamma_{FR} < 0$: that if past change has been in one direction then future change over any forecast horizon will on the average be in the other direction, and that, moreover, the larger the magnitude of the past change the more pronounced the typical reversal.

These two cases provide the basis for an alternate definition of the model developed in section 3.1. Construct the following weighted average of the definitions of each special case stated in (3.3.1) and (3.3.2):

$$\bar{\bar{X}}_t = \beta X_{t-1} + (1-\beta)\ \bar{\bar{X}}_{t-1}, \quad 0 \leqq \beta \leqq 1. \tag{3.3.5}$$

It is evident that the model defined by (3.3.5) is identical to (3.1.8) for $\beta = \lambda$. The "proportional effect" model or exponentially-weighted moving average can thus be regarded as in effect postulating that a time series is a linear combination of a random walk and of a set of independent samplings from a common probability distribution.

3.4 Implications of λ for the variance of X

An analysis of the stability of the variance of X through time can provide a reasonably fruitful test of the applicability of either a random walk or a set of independent identically-distributed observations as a description of X, and in the process provide an estimate of λ. From (3.1.2) and the postulated

independence of successive ζ_j, it follows that

(3.4.1) $$\text{var} \quad X_t = (t\,\lambda^2 + 1) \quad \text{var}\ \xi$$

Similarly,

(3.4.2) $$\text{var} \quad X_{t-R} = [(t-R)\,\lambda^2 + 1] \quad \text{var}\ \xi$$

so that

(3.4.3) $$\text{var} \quad X_t = \text{var}\ X_{t-R} \left[\frac{t\,\lambda^2 + 1}{(t-R)\lambda^2 + 1}\right]$$

If $\lambda = 0$, as in the case of a time series formed by successive independent, identically distributed observations, then the variance of X is unchanged over time. If λ 0, then var (X) will increase, though at a declining rate. If $\lambda = 1$, as in the case of a random walk, then (3.4.3) degenerates to

(3.4.4) $$\frac{\text{var}\ X_t}{\text{var}\ X_{t-R}} = \frac{t + 1}{t-R+1} \quad ,$$

a well-known property of a random walk.

Alternatively, attention may be focused on the ratio of the variance of Δ_R to that of Δ_1, where $\Delta_1 = X_t - X_{t-1}$. This ratio (which will be denoted by J_R) provides a more sensitive test of whether a time series is a random walk. From (3.2.3), it is evident that

(3.4.5) $$\text{var} \quad \Delta_R = [\lambda^2 R + 2(1-\lambda)] \quad \text{var}\ \zeta$$

so that, for $R > 0$,

(3.4.6) $$J_R = \frac{\text{var}\ \Delta_R}{\text{var}\ \Delta_1} = \frac{\lambda^2 R + 2(1-\lambda)}{\lambda^2 + 2\,(1-\lambda)}$$

For a random walk, $\lambda = 1$ so that $J_R = R$. The variance of change in X is thus then

proportional to the length of time over which such change is analyzed. For the "timeless" time series defined by the second special case, $\lambda = 0$ and the variance of change in X is unrelated to the duration of change. More generally, if $0 \leq \lambda \leq 1$, then the relationship between J_R and R is shown by Chart 2 for different values of λ.

Paul Cootner [15] has found that, for stock prices, the variance of price changes over periods of R weeks was generally significantly different from R times the variance of changes over one week.[1] It is thus evident that the time-pattern of stock market prices is not a random walk (as hypothesized in the studies by Alexander [1] and Kendall [35] cited <u>supra</u> as applications of a random walk). In addition, he found significant curvature in the relation between J_R and R. While it may be that this evidence substantiates either the model defined in section 3.1 or the constrained random walk suggested by Cootner as a model of stock prices,[1] neither is sufficient in itself to account for changes in dJ/dR. It may be that such changes can be explained in terms of a drift with time in the reflecting barriers of Cootner's model or in the expected value of ζ_t in the model of section 3.1. It is in any case evident that the simple model defined in section 3.1 is not adequate as there postulated.

I shall not pay further attention in this paper to the tests of observed time series suggested by this discussions. A discussion of J(R) for the other models advanced in this paper will be presented elsewhere.

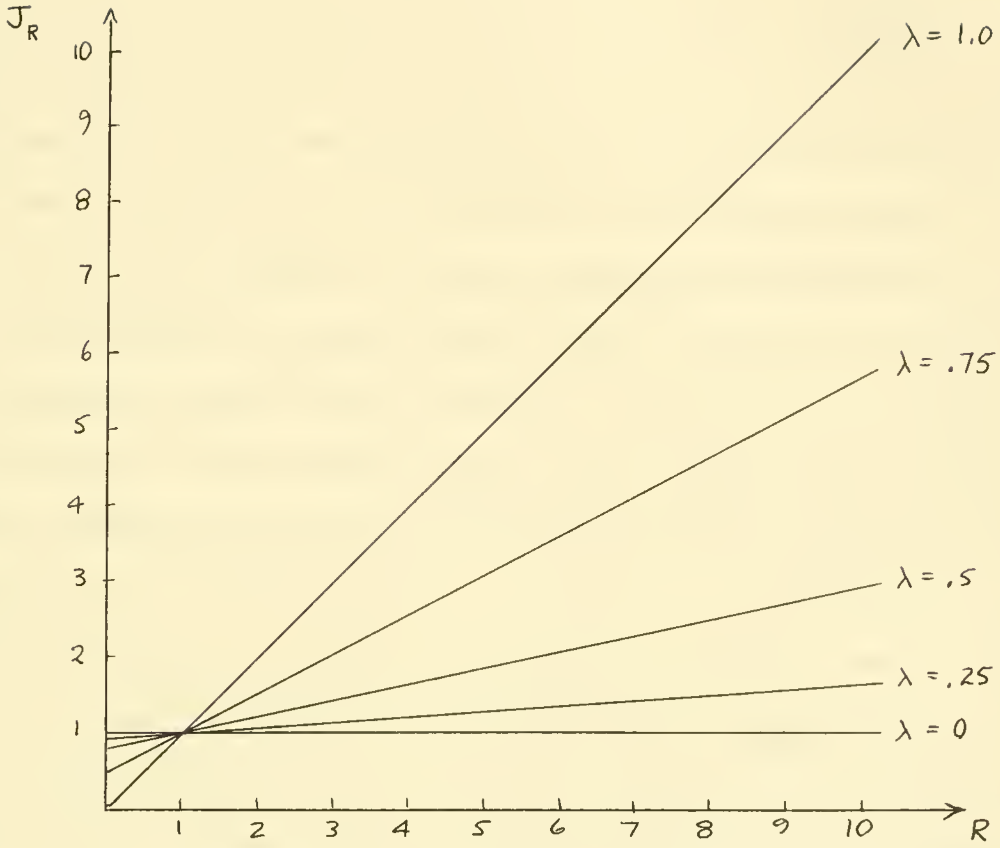
---

[1] The model defined in section 3.1 can be considered as having essentially identical properties to those of a random walk constrained by a large number of partly reflecting barriers defined at a large number of distances from some "normal" price defined in each period. (A physical analog would be a set of successive lattices of crystals, in which each lattice reflects only a portion of the particles which have passed through previous lattices). In terms of a model of common stock prices, the model defined in section 3.1 thus assumes a spectrum of classes of investors in place of the two classes postulated by Cootner, each class in the spectrum having somewhat different opportunity costs of search and/or different potential amounts of information which might be unearthed by such search.)

CHART 2

RELATION BETWEEN J AND R

FOR SELECTED VALUES OF λ



4. A two-stage process

The simple exponentially-weighted moving average discussed in section 3 is a descriptive model of an economic time series only if the following assumptions are satisfied: (1) each observation of the time series is determined up to a current random shock by the effects of previous random shocks, (2) each random shock is a realization of a pure random Gaussian process with zero mean, and (3) the effect of each random shock is all identical for all subsequent time. In this section I shall relax assumption (2) while maintaining assumptions (1) and (3). To do so I shall define a two-stage model of X(t), assuming its second stage to be the "proportional effect" model developed in section 3.1. The output of the second-stage model is thus defined as the sum of a current random shock plus the persistent components of previous random shocks, each shock being in turn defined as the output of a first-stage model. In section 4.1, I shall first of all specify this first-stage model as a linear transformation of the pure random Gaussian process $\zeta(t)$, relaxing merely the assumption that the random shocks are a realization of a process with zero mean. Even maintaining all other assumptions previously specified about the random shocks, the model of section 4.1 necessitates a substantial modification of the conclusions reached in section 3 regarding the behavior of γ_{FR}. I shall then go on in section 4.2 to relax the assumption that each random shock is identically distributed and in section 4.3 to relax the assumption that each shock is independently distributed.

4.1 Random shocks with non-zero mean

Other than to assume that a random shock is a realization of the pure random process $\left\{\xi(t)\right\} = b \left\{\zeta(t)\right\}$, as postulated in equation (3.1.1), the next simplest assumption that can be made is to continue to assume that it is a

realization of a pure random process (so that the elements of any set of random shocks are independently and identically distributed) but at the same time to relax the assumption that the mean of the process is zero. Define a process $\{\xi(j)\}$ such that

(4.1.1) $$\xi_j = a + b\zeta_j$$

The expected value of ξ_j is then $\underline{a}$, and $\underline{b}^2$ is the variance of the process. Since ξ_j is a linear transformation os a pure random Gaussian process, it is itself a pure random Gaussian process.[1] The only assumption governing ξ_j which has been relaxed is that $\{\xi_j\}$ be a linear homogeneous transformation of $\{\zeta_j\}$.

The implications of this relaxation for the nature of the stochastic process $X(t)$ of which an observed time series is a realization can be shown quickly. Applying the "second-stage" model of section 3 to the random shocks defined by (4.1.1).

(4.1.2) $$X_t = \xi_t + \lambda \sum_{j<t} \xi_j$$

as in (3.1.2). Substituting (4.1.1) in (4.1.2), the two stages of the model can be combined to define X_t in terms of the underlying pure random process $\{\zeta_j\}$:

(4.1.3) $$X_t = a(1 + \lambda t) + b\left[\zeta_t + \lambda \sum_{j<t} \zeta_j\right]$$

so that $\bar{\bar{X}}_t$, the conditional expectation of X_t given previous realized values of $\{\zeta_j\}$, is

(4.1.4) $$\bar{\bar{X}}_t = a(1 + \lambda t) + b\lambda \sum_{j<t} \zeta_j .$$

---

1 It may in some circumstances also be appropriate to relax the assumption that $\{\xi(j)\}$ is not only a pure random process but also Gaussian -- in other words, to relax the restrictions on third and higher moments of the distribution of each ξ_j. This extension will not be made in this paper.

It is evident that equation (4.1.4) does not reduce to an exponentially-weighted average of previous observed values of X. From (4.1.3) and (4.1.4), $X_t - \bar{\bar{X}}_t = b\,\xi_t$ and $\bar{\bar{X}}_t - \bar{\bar{X}}_{t-1} = \lambda\,\xi_{t-1} = \lambda\,[a + b\,\xi_{t-1}]$. Consequently,

(4.1.5) $$\bar{\bar{X}}_t - \bar{\bar{X}}_{t-1} = \lambda\, a + \lambda\,[X_{t-1} - \bar{\bar{X}}_{t-1}]$$

so that, rearranging and iterating,

(4.1.6) $$\bar{\bar{X}}_t = \lambda \sum_{j=0}^{t-1} (1-\lambda)^{t-j-1} X_j + a\,[1 - (1-\lambda)^t].$$

This expression diverges from the exponentially-weighted average of (3.1.9) by an amount which (for t greater than say $10/\lambda$), can be effectively regarded as a constant equal to a, the mean value of ξ_t.

The implications of this two-stage process for γ_{FR} likewise diverge from those of the model discussed in section 3. First of all, it is apparent from (4.1.3) that

(4.1.7) $$\Delta_F = a\lambda F + b\,[\xi_{t+F} + \lambda \sum_{j=t+1}^{t+f-1} \xi_j - (1-\lambda)\,\xi_t]$$

so that the unconditional expected value of Δ_F is $a\lambda F$ rather than zero as in the single-stage model. Because of the trend thus introduced, γ_{FR} is a function of F as well as of λ and R. In addition, it is evident that γ_{FR} is also a function of the parameters of the first-stage model defined in (4.1.1). From (4.1.3),

(4.1.8) $$\Delta_R = a\lambda R + b[\xi_t + \lambda \sum_{j=t+1}^{t+F-1} \xi_j - (1-\lambda)\,\xi_{t-R}]$$

Substituting (4.1.7) and (4.1.8) into (2.3), and forming expected values of the resultant product moments, we obtain

(4.1.9) $$\gamma_{FR} = \frac{\lambda^2 a^2 \; FR + b^2 \; (\lambda-1)}{\lambda^2 a^2 R^2 + b^2 \; [\lambda^2 R + 2 \; (1-\lambda)]}$$

which reduces to (3.2.6) for a = 0. It is evident that γ_{FR} is in this model not necessarily non-positive. Indeed, since the denominator is positive, γ_{FR} is positive if and only if the following inequality is satisfied:

(4.1.10) $$\lambda^2 a^2 \; FR \qquad b^2 \; (1-\lambda).$$

Condition (4.1.10) can be expressed in the following way: γ_{FR} is positive if λ is greater than some λ_o defined by the equation

(4.1.11) $$\frac{\lambda_o^{\,2} \; FR}{1 - \lambda_o} = \frac{b}{a} = V \;,$$

where V is the coefficient of variation of . Chart 3 shows the relationship between λ_o and V for selected values of the product of F and R.

As is intuitively evident from the fact that $\lambda^2 a^2 FR$, the left term of expression (4.1.10), is the product of the trends defined over the lengths of the forecast and record horizons, chart 3 shows that the longer these horizons (other things being equal), the greater the range of λ and V for which γ is positive. The chart also shows, however, that the size of the range of FR and λ for which γ_{FR} is positive decreases as V increases. Indeed, for infinite V (as in the case of a = 0 for any non-zero b), γ_{FR} is non-positive for all defined values of λ and is strictly negative if λ 1.

For any values of λ, F, and R there is thus some V' such that, for V V', γ_{FR} is negative. Defining the degree of regressivity in a time series as the degree to which γ_{FR} is negative,[1] it is evident that the degree of regressivity

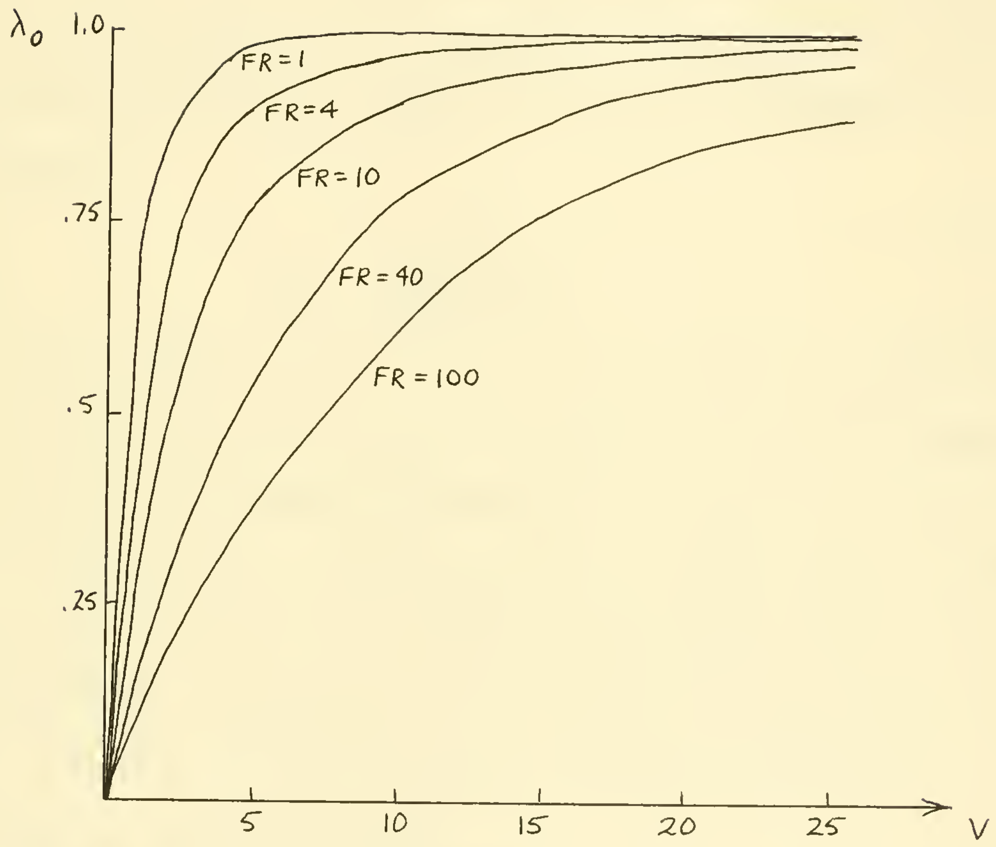
---

[1] This follows the definition of regressivity used [9, p. 246] and [10].

CHART 3

RELATIONSHIP BETWEEN λ_0 AND V

FOR SELECTED FR



is an increasing function of V but a decreasing function of λ and FR. Holding V and FR constant, the degree of regressivity in a time series is a decreasing function of λ alone, as in the model discussed in section 3.2.

It may be useful to discuss the implications of the "first-stage" model defined in (4.1.1) for the values of γ_{FR} in the two special cases of the "second-stage" model of section 3 defined by $\lambda = 0$ and $\lambda = 1$. If $\lambda = 0$, then, substituting in (4.1.9), $\gamma_{FR} = -1/2$. In this special case the properties of γ_{FR} are thus independent of the parameters of the "first-stage" model. This is not true for the random walk defined by $\lambda = 1$, however. In this case substitution of $\lambda = 1$ into (4.1.9) yields

(4.1.12)
$$\gamma_{FR} = \frac{a^2 FR}{a^2 R^2 + b^2 R}$$

which is non-positive only if $a = 0$. This result, again, is consonant with what one should intuitively expect. Since a random walk is merely a summation of previous random shocks, a non-zero expected value of those shocks results in the random walk being defined around some positive or negative trend. Consequently $E\, [X_{t+F} \;\; X_t] - X_t = aF$ so that, since Δ_R on the average equals aR, the directions of change tend to be the same over both forecast and record horizons -- a statement which is equivalent to stating that $\gamma_{FR} \;\; 0$ for non-zero values of a.

4.2 Non-identically distributed shocks

Up to this point, it has been assumed that the random shocks described by the intermediate process (j) are independently and identically distributed. I shall now drop the assumption that they are identically distributed, first examining the case where (j) is merely heteroscedastic, and then going on to examine the more general case where not only Var (j) but also E [(j)] are dependent on time.

4.2.1. Heterscedastic $\xi(j)$ I shall assume time-dependence in the variance of $\xi(j)$ such that Var $\xi_j = (1 + f)$ Var ξ_{j-1} for all j, where $f > -1$. Further defining Var $\xi_0 = b^2$ Var $\zeta_0 = b^2$, then, since Var $\zeta_j =$ Var ζ_0,

(4.2.1) $$\text{Var } \xi_j = b^2 (1+f)^j \text{ Var } \zeta_j .$$

Consequently, defining as before $E(\xi_j) = a$,

(4.2.2) $$\xi_j = a + b\sqrt{(1+f)^j}\,\zeta_j .$$

Substituting this definition of ξ_j in equation (4.1.2), the combined two-stage process thus becomes

(4.2.3) $$X_t = a(1+\lambda t) + b[\zeta_t\sqrt{(1+f)^t} + \lambda \sum_{j<t} \zeta_j \sqrt{(1+f)^j}] ,$$

which reduces to equation (4.1.3) when f = 0. It can easily be shown that

(4.2.4) $$\gamma_{FR} = \frac{\lambda^2 a^2 FR + b^2(\lambda-1)(1+f)^t}{\lambda^2 a^2 R^2 + b^2[S]}$$

$$\text{where } S = (1+f)^t + \lambda^2 \sum_{j=t-R}^{t-1} (1+f)^j + (1-2\lambda)(1+f)^{t-R}$$

This equation of course reduces to equation (4.1.9) if f = 0. It is evident that the degree of regressivity in X is related to f. Specifically, it is clear that γ_{FR} is positive only if λ is greater than some λ_0^*, where λ_0^* is larger than the λ_0 defined in equation (4.1.11) if $f > 0$ but smaller than λ_0 if $f < 0$. As in section 4.1, γ_{FR} is non-positive for all defined values of λ and f if a = 0.

4.2.2. Time dependent mean: Case I If ξ not only is heteroscedastic but also has an expected value which drifts with time, it is necessary to further modify earlier statements about the behavior of γ_{FR}. I shall examine γ_{FR} for two different models of time-drift in ξ: assuming in the first case that $E(\xi_t) = a + ct$

and then in the second case that $E(\xi_t) = a + d \log (t+1)$. Heteroscedasticity will be defined as in section 4.2.1.

In Case I, it is assumed that there is a constant linear drift in ξ so that $E(\xi_j - \xi_{j-1}) = c$. Defining $E(\xi_0) = a$, then $E(\xi_j) = a + cj$, so that, from this definition and from (4.2.1),

$$(4.2.5) \qquad \xi_j = a + cj + b\sqrt{(1+f)^j}\,\zeta_j\,.$$

Substituting (4.2.5) in equation (4.1.2),

$$(4.2.6) \qquad X_t = a(1+\lambda t) + ct[1 + \frac{\lambda}{2}(t-1)] + b\sqrt{(1+f)^t}\,\zeta_t + \lambda b \sum_{j<t} \sqrt{(1+f)^j}\,\zeta_j$$

Consequently, deriving γ_{FR} as before,

$$(4.2.7) \qquad \gamma_{FR} = \frac{FR[L] + b^2(\lambda-1)(1+f)^t}{b^2[S] + R^2[K]}$$

where
$$L = a^2\lambda^2 + \lambda ac[2(1+\lambda t) + \lambda(\frac{F-R}{2} - 1)] + c^2[1 + \lambda(t + \frac{F-1}{2})][1 + \lambda(t - \frac{R+1}{2})]$$

$$S = (1+f)^t + \lambda^2 \sum_{j=t-R}^{t-1} (1+f)^j + (1-2\lambda)(1+f)^{t-R}$$

$$K = a^2\lambda^2 + 2\lambda ac\,[1 + \lambda(t - \frac{R+1}{2})] + c^2[1 + \lambda(t - \frac{R+1}{2})]$$

If $c = 0$, equation (4.2.6) reduces to (4.2.3) and equation (4.2.7) reduces to (4.2.4). As before, γ_{FR} is positive if and only if the numerator of (4.2.7) is positive. Consequently γ_{FR} is positive if and only if

$$(4.2.8) \qquad \lambda^2 a^2 FR + cFR[W] > b^2 (1-\lambda)(1+f)^t$$

$$\text{where} \quad W = (F-R) \frac{\lambda}{2} [\lambda a + c + c\lambda - \frac{\lambda}{2}] - \frac{\lambda^2}{4} FR$$

$$+ t\lambda^2 [2a + \frac{2c}{\lambda} + c(t-1)]$$

$$+ \lambda (2a - \lambda a - c + \frac{\lambda}{4}) + c$$

This condition can be written in the following way: γ_{FR} is positive if and only if λ is greater than some λ_0^{**} defined by the equation

$$(4.2.9) \qquad (\lambda_0^{**})^2 FR - \frac{c}{2a^2} FR [W] = \frac{b^2}{a^2} (1+f)^t (1-\lambda_0^{**})$$

where W is defined in (4.2.8).

Comparing this condition to the one discussed in subsection 4.2.1 for series which are formed from random shocks that are time-dependent only in their variances, it is clear that of course $\lambda_0^{**} = \lambda_0^*$ if $c = 0$. For non-zero values of c the comparison is not so clear, however, for λ_0^{**} may be larger or smaller than λ_0^* depending on the values of a and c and of the time-parameters F, R, and t. The reason for this is apparent enough: in the general case, it cannot be specified whether the time-dependent and non-time-dependent trends defined respectively by c and a reinforce one another or the opposite. If they reinforce one another, the effect will be to make γ_{FR} larger than it would be in the absence of the time-dependent trend. If they counter each other, the marginal effect of the time-dependent trend is to increase the regressivity of the series.

It is possible to make more precise general statements in the special cases defined by $\lambda = 0$ and $\lambda = 1$. If $\lambda = 0$, then, substituting this value of X in (4.2.7),

$$\gamma_{FR} = \frac{FR\ c^2 - b^2\ (1+f)^t}{b^2[(1+f)^t + (1+f)^{t-R}] + R^2c^2} \tag{4.2.10}$$

It is thus evident that γ_{FR} is equal to - 1/2 for all values of a, b, t, F, and R only if c = f = 0. If only c = 0, then γ_{FR} is still strictly negative, but can take on any values in the open interval between - 1 and zero as indicated by

$$\gamma_{FR} = - \frac{(1+f)^R}{(1+f)^R + 1} \tag{4.2.11}$$

If c is also non-zero, it is evident from (4.2.10) that γ_{FR} may even be positive. Time dependence in the random shocks thus substantially affects the behavior of γ_{FR} in this case.

Turning to the second boundary-defined special case, equation (4.2.7) reduces to the following expression when $\lambda = 1$:

$$\gamma_{FR} = \frac{FR \left\{a^2 + ac\left[2(1+t) + \frac{F-R}{2} - 1)\right] + c^2[1+t+\frac{F-1}{2}][1+t-\frac{R+1}{2}]\right\}}{b^2 \sum_{j=t-R+1}^{t} (1+f)^j + R^2[a^2 + 2ac\ (1+t-\frac{R+1}{2}) + c^2(1+t-\frac{R+1}{2})^2]} \tag{4.2.12}$$

The most obvious point that can be made about the range of γ_{FR} defined by this equation is that of course it reduces to that defined by (4.1.12) if c = f = 0. It is clear that γ_{FR} is not necessarily non-positive if a = 0. If a = 0, then (4.2.12) reduces to

$$\gamma_{FR} = \frac{FR\ c^2(1+t+\frac{F-1}{2})(1+t - \frac{R+1}{2})}{b^2 \sum_{j=t-R+1}^{t} (1+f)^j + c^2R^2(1+t-\frac{R+1}{2})^2} \tag{4.2.13}$$

so that γ_{FR} is in this case non-positive if and only if

$$FR\ c^2(1+t+\frac{F-1}{2})(1+t-\frac{R+1}{2}) \leqq 0. \tag{4.2.14}$$

Since $t > R$, this condition is satisfied only if $c = 0$. γ_{FR} can thus be positive even when $a = 0$, as one might expect from the fact that any trend around which the random walk is defined has the same sign in both Δ_F and Δ_R no matter whether or not the trend is time-dependent.

4.2.3 Time-dependent mean: Case II It may for some series be more realistic to assume the existence of a decreasingly-important drift in ξ rather than the constant drift postulated in subsection 4.2.2. One model of drift which satisfies the assumption of a constantly decreasing expected value of change in ξ is that obtained by specifying that $E[\xi_j - \xi_{j-1}] = d \log[(j+1)/j]$. Specifying this model and further specifying as before $E(\xi_0) = a$ and heteroscedasticity as in (4.2.1), then $E(\xi_j) = a + d \log(j+1)$ and

$$\xi_j = a + d \log(j+1) + b\sqrt{(1+f)^j}\,\zeta_j \tag{4.2.15}$$

Substituting this equation in (4.1.2),

$$\begin{aligned} X_t &= a(1+\lambda t) + d[\log(t+1) + \lambda \sum_{j<t} \log(j+1)] \\ &\quad + b\sqrt{(1+f)^t}\,\zeta_t + \lambda b \sum_{j<t} \sqrt{(1+f)^j}\,\zeta_j\,. \end{aligned} \tag{4.2.16}$$

Comparing this model with that defined in (4.2.6), it is evident that the only difference between the two models is the lesser importance of the time-dependent trend in (4.2.16). The behavior of γ_{FR} will consequently correspond to that described in subsection 4.2.2, subject to this qualification.

4.3. Interdependent shocks

To generalize the first-stage model thus far developed to allow for the possibility of interdependency in the random shocks generated as inputs for the second-stage model, I assume an additive difference between the unconditional

expected value of ξ_j, $E(\xi_j)$, and the expected value of ξ_j given the values of preceding shocks, $E(\xi_j | \xi_{j-1}, \xi_{j-2}, \ldots)$. I shall postulate as in section 4.2.2 that $E(\xi_j) = a + c_j$, and shall further postulate the first order dependence in ξ defined for $j > 0$ by

$$(4.3.1) \qquad E(\xi_j | \xi_{j-1}, \xi_{j-2}, \ldots) = a + cj + g[\xi_{j-1} - E(\xi_{j-1})].$$

If $g = 0$, this conditional expected value reduces to the unconditional expectation $E(\xi_j)$. If $g < 0$, then the series formed by the deviations of the shocks from their trend is regressive; if $g > 0$, the opposite is true. An interesting case is defined by $g = 1$; in this case ξ_j is defined as a random walk around the trend defined by the parameters a cna c since (4.3.1) then reduces to

$$(4.3.2) \qquad E(\xi_j | \xi_{j-1}, \xi_{j-2}, \ldots) = \xi_{j-1} + c.$$

Further postulating that the variance of ξ_j is defined by (4.2.1) if $g = 0$, it is evident from (4.3.1) that

$$(4.3.3) \qquad \xi_j = a + cj + b\sqrt{(1+f)^j}\,\zeta_j + g[\xi_{j-1} - a - c(j-1)]$$

for $j > 0$ so that, iterating within the right-hand term of (4.3.3),

$$(4.3.4) \qquad \xi_j = a + cj + b\sum_{i=0}^{j} g^i \sqrt{(1+f)^{j-i}}\,\zeta_{j-i}$$

for all j. It is of course possible to assume more complicated interdependence in the ξ, but I shall restrict myself in this paper to that defined by (4.3.1).

Substituting (4.3.4) in the second-stage model defined by (4.1.2), the combined two-stage model thus becomes

$$(4.3.5) \qquad X_t = a + ct + b \sum_{i=0}^{t} g^i \sqrt{(1+f)^{t-i}}\, \zeta_{t-i}$$

$$+ \lambda \sum_{j=0}^{t-1} \left[a + cj + b \sum_{i=0}^{j} g^i \sqrt{(1+f)^{j-i}}\, \zeta_{j-i} \right]$$

Since

$$(4.3.6) \qquad \sum_{j=0}^{t-1} \sum_{i=0}^{j} g^i \sqrt{(1+f)^{j-i}}\, \zeta_{j-i} = \sum_{j=0}^{t-1} \sum_{i=0}^{t-j} g^i \sqrt{(1+f)^{j}}\, \zeta_j$$

we can, after closing sums, rewrite (4.3.5) as

$$(4.3.7) \qquad X_t = a(1+\lambda t) + ct\,[1 + \tfrac{\lambda}{2}(t-1)] + b \sqrt{1+f)^t}\, \zeta_t$$

$$+ b \sum_{j=0}^{t-1} [g^{t-j} + \frac{\lambda}{1-g}(1+g^{t-j+1})] \sqrt{1+f)^j}\, \zeta_j.$$

Consequently, proceeding in the usual manner to find γ_{FR},

$$(4.3.8) \qquad \gamma_{FR} = \frac{FR\,[L] + b^2[Z]}{b^2\,[V] + R^2[k]} -$$

where $L = a^2\lambda^2 + \lambda ac\,[2(1+\lambda t) + \lambda\,(\frac{F-R}{2} - 1)]$

$$+ c^2\,[1 + \lambda\,(t + \frac{F-1}{2})][1 + \lambda\,(t - \frac{R+1}{2})]$$

$$Z = [g^F + \frac{\lambda}{1-g}(1+g^{F+1}) - 1]\,(1+f)^t$$

$$+ \sum_{j=t-R+1}^{t-1} [g^{t-j} + \frac{\lambda}{1-g}(1+g^{t-j+1})][(g^{t+F-j} - g^{t-j})(1 + \frac{\lambda g}{1-g})]$$

$$+ [g^R + \frac{\lambda}{1-g}(1+g^{R+1}) - 1][(g^{F+R} - g^R)(1 + \frac{\lambda g}{1-g})](1+f)$$

$$+ \sum_{j=0}^{t-R-1} (g^{t-j} - g^{t-R-j})(g^{t+F-j} - g^{t-j})(1 + \frac{\lambda g}{1-g})^2\,(1+f)^j$$

$$K = a^2\lambda^2 + 2\lambda ac\left(1+\lambda\left[t - \frac{R+1}{2}\right]\right) + c^2\left(1+\lambda\left[t - \frac{R+1}{2}\right]\right)$$

$$V = (1+f)^t + \sum_{j=t-R+1}^{t-1}\left[g^{t-j} + \frac{\lambda}{1-g}\left(1+g^{t-j+1}\right)\right]^2 (1+f)^j$$

$$+ \left[g^R + \frac{\lambda}{1-g}\left(1+g^{R+1}\right) - 1\right]^2 (1+f)^{t-R}$$

$$+ \sum_{j=0}^{t-R-1}\left[\left(g^{t-j} - g^{t-R-j}\right)\left(1 + \frac{\lambda g}{1-g}\right)\right]^2 (1+f)^j$$

The equation is somewhat formidable in appearance. Nevertheless, it is different from equation (4.2.7) only in the expressions Z and V, which are themselves functions only of f, g, λ, and the time parameters F, R, and t. The behavior of γ_{FR} with respect to changes in a and c is thus qualitatively no different from that described in sections 3.1 and 3.2.

The effect of the interdependence introduced by non-zero values of g can be pinpointed by examining the behavior of γ_{FR} with respect to g in the special case where no trend is posulated in ξ_j. If a = c = 0, then (4.3.8) reduces to γ_{FR} = Z/V, where Z and V are defined as above.

5. An alternate second-stage model

Up to this point, it has been assumed that the effects of random influences occurring in a given period can be dichotomized into two parts, consisting of those effects which persist beyond the given period and those which do not. The persistent effects of random influences are defined to be truly persistent, in that a random shock is assumed to have the same effect on the realized variable for all subsequent time.

Since it may be desirable to regard the time-pattern of the effects of a random shock as other than the simple dyadic specification thus far discussed, an alternate second stage model is analyzed in this section. The model itself is presented in section 5.1. The behavior of γ_{FR} given this model is then analyzed in section 5.2 and in section 5.3.

5.1 Pascal-distributed effects

The essence of the second stage of the two-stage model which has been presented in this paper is a specification of the nature of the effect of a random shock occurring in any period upon the level in that and other periods of the variable being analyzed. (The stochastic process generating that random shock is specified by the first-stage model). To put this statement another way, let me use the notation $F_{j+k}(\xi_j)$ to denote the effect in period $j + k$ of a random shock ξ_j occurring in period j. The model specified in section 3 can then be formulated in the following way:

(5.1.1)
$$\begin{aligned} F_{j+k}(\xi_j) &= \xi_j\,, \quad k = 0 \\ &= \lambda\,\xi_j\,, \quad k = 0, \quad 0 \leqq \lambda \leqq 1. \end{aligned}$$

In many circumstances it may seem inappropriate to assume that the effect of a random shock is identical in all subsequent periods or -- the same thing -- that the sum of the effects of a random shock on all subsequent periods is non-finite. One alternative is to postulate that the effect of a random shock continuously declines as time goes on. A convenient specification of this alternative is the exponential decay defined by

$$F_{j+k}(\xi_j) = \delta F_{j+k-1}(\xi_j), \quad k > 0, \quad 0 \leqq \delta < 1. \tag{5.1.2}$$

Specifying $F_j(\xi_j) = (1-\delta)\xi_j$, it follows from (5.1.2) that

$$F_{j+k}(\xi_j) = (1-\delta)\delta^k \xi_j \tag{5.1.3}$$

for $k > 0$.

The model described by this equation is similar to that defined by (4.1.1) in that the special cases of each defined by $\delta = 0$ and $\lambda = 0$ are identical. In other respects it is quite dissimilar. In particular, it can be easily shown that, since $\delta < 1$,

$$\lim_{n \to \infty} \sum_{k=0}^{n} F_{j+k}(\xi_j) = \xi_j \tag{5.1.4}$$

so that there is a finite upper bound on the sum of all effects of the shock. The model expressed by (5.1.2) can thus be viewed as an allocation over subsequent time of the random shock ξ_j.

The operational definition of ξ_j implied by this model is substantially different from that implied by the proportional effect model defined in (5.1.1). In that model, the random shock ξ_j was defined as its effect in period j. By contrast, ξ_j is defined in the exponential decay model as its total effect in all periods; its effect in period j is merely $(1-\delta)\xi_j$. These two definitions

coincide only for the special cases where δ and λ equal zero. As δ increases the definitions diverge. Indeed,

$$(5.1.5) \qquad \lim_{\delta \to 1} F_j(\xi_j) = \lim_{\delta \to 1} F_{j+k}(\xi_j) = 0$$

so that the definitions would infinitely diverge if δ were permitted to equal unity. The fact that $F_j(\xi_j)$ asymtotically vanishes as δ approaches unity merely means that the ratio of the total effect of ξ_j to its effect in period j increases indefinitely as δ goes to unity -- or, to put it differently, as the observed time series to be defined by (5.1.3) approaches a random walk.

As this statement implies, the structure of the time series defined by an exponential decay model approaches that of the special case of (5.1.1) defined by $\lambda = 1$, as δ approaches unity. However, this convergence is only readily apparent for values of δ close to unity. The difference between the time pattern of the effects of ξ_j specified by (5.1.1) and that specified by (5.1.3) may be defined as the difference between the values of the ratio $F_{j+k}(\xi_j)/F_j(\xi_j)$ in each model. Chart 4 shows values of this difference corresponding to various values of λ and δ. Since this difference is simply the value of $(\lambda - \delta^k)$, it is evident that this difference increases virtually until the "random walk" special case is reached.

The exponential decay model specified by (5.1.2) is interesting among other reasons because it reduces to a Markov process.[1] Defining X_t, as before, as the sum of a current random shock and the persistent components of previous shocks,

$$(5.1.6) \qquad X_t = (1-\delta)\,\xi_t + \sum_{j=0}^{t-1} (1-\delta)^{t-j}\,\xi_j$$

so that

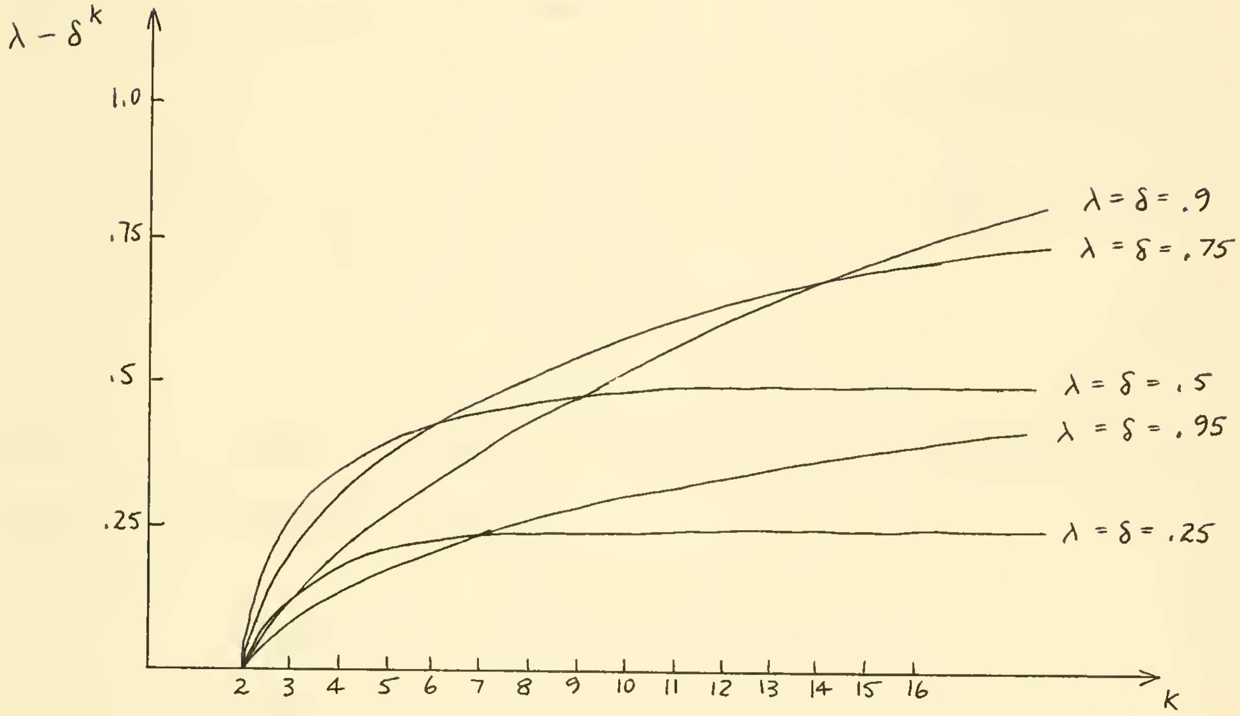
$$(5.1.7) \qquad X_t = \delta X_{t-1} + \xi_t (1-\delta)$$

[1] There is a voluminous literature on Markov processes. For surveys of the statistical theory, see Billingsley [6][7] and Feller [22 , chapters 15 and 16]. Some economic applications (to marketing problems) are discussed in Harary and Lipstein [31], Heiniter and Magee [32], and Maffei [40].

CHART 4

DIFFERENCES BETWEEN VALUES OF $F_{j+k}(\xi_j)/F_j(\xi_j)$

FOR "δ" AND "λ" MODELS



The observed time series is thus by this model a first-order Markov process so long as the random shocks ξ_j are independently and identically distributed.

In a number of applications, it is unsatisfactory to have to assume that the greatest effect of a random shock is at the time of its occurrence. The effect of random shocks may take time to build up to a maximum before beginning to tail off. Consequently it would be useful to specify a more general model of the time pattern of random shocks' effects.

As Robert Solow [58] has pointed out, a natural generalization of (5.1.3) exists by virtue of the fact that the successive ratios $(1/\xi_j)[F_{j+k}(\xi_j)]$ are terms of the geometric probability distribution, a special case of the Pascal distribution.[1] Consequently we may write

$$F_{j+k}(\xi_j) = \binom{r+k-1}{k} (1-\delta)^r \delta^k \xi_j \tag{5.1.8}$$

where r is a positive integer. This expression reduces to (5.1.3) for r = 1. Since, from the binomial theorem,

$$\lim_{n\to\infty} \sum_{k=0}^{n} \binom{r+k-1}{k} \delta^k = (1-\delta)^{-r}, \tag{5.1.9}$$

it is evident that

$$\lim_{n\to\infty} \sum_{k=0}^{n} F_{j+k}(\xi_j) = \xi_j \tag{5.1.10}$$

as in (5.1.4). The Pascal distribution thus provides a general specification of the distribution of the total effects of each random shock over subsequent time.

The difference between the Pascal distribution defined by different values

---

[1] The Pascal distribution is discussed in Solow [64, pp. 394-396] and Feller [22, pp. 155-157]. This distribution can be regarded as the result of r exponential decay model cascaded in series; it is so discussed in Forrester [24, chapter 9]. See also Theil and Stern [68, pp. 113-115].

of r is illustrated in chart 5, which presents an indication of the time pattern of effects associated with various values of r.[1] For convenience in presentation $F_{j+k}(\xi_j)$ is defined as a continuous function of k. The axes of chart 5 are defined so as to collapse the family of distributions defined by different values of δ into a simple distribution for each given value of r.

As chart 5 indicates, the maximum effect per period of a random shock moves outward as r is increased. For r = 1, the maximum effect of a shock is in the period in which the shock occurs. For r = 2, the immediate effect of a random shock is zero, at which time the rate of change of $F_{j+k}(\xi_j)$ is at a maximum.[2] For $r > 2$, both the immediate effect of the shock and its rate of change at that time are zero. There seems little point in postulating $r > 3$ for most economic applications, since the qualitative nature of the time-pattern of delayed effects is specified by any $r > 2$. For larger, the resultant distribution becomes both less skewed and more leptokurtic: the length of time before $F_{j+k}(\xi_j)$ reaches a given value increases (though at a decreasing rate), the rate of increase once a significant level of $F_{j+k}(\xi_j)$ is reached becomes larger, and the rate of approach to zero of $F_{j+k}(\xi_j)$ is faster once the model value of the distribution is passed.

Given the specification of $F_{j+k}(\xi_j)$ defined in (5.1.8), the second-stage model can be quickly obtained as the sum of the currently persistent components (i.e., the current effects in period t) of both current and previous

---

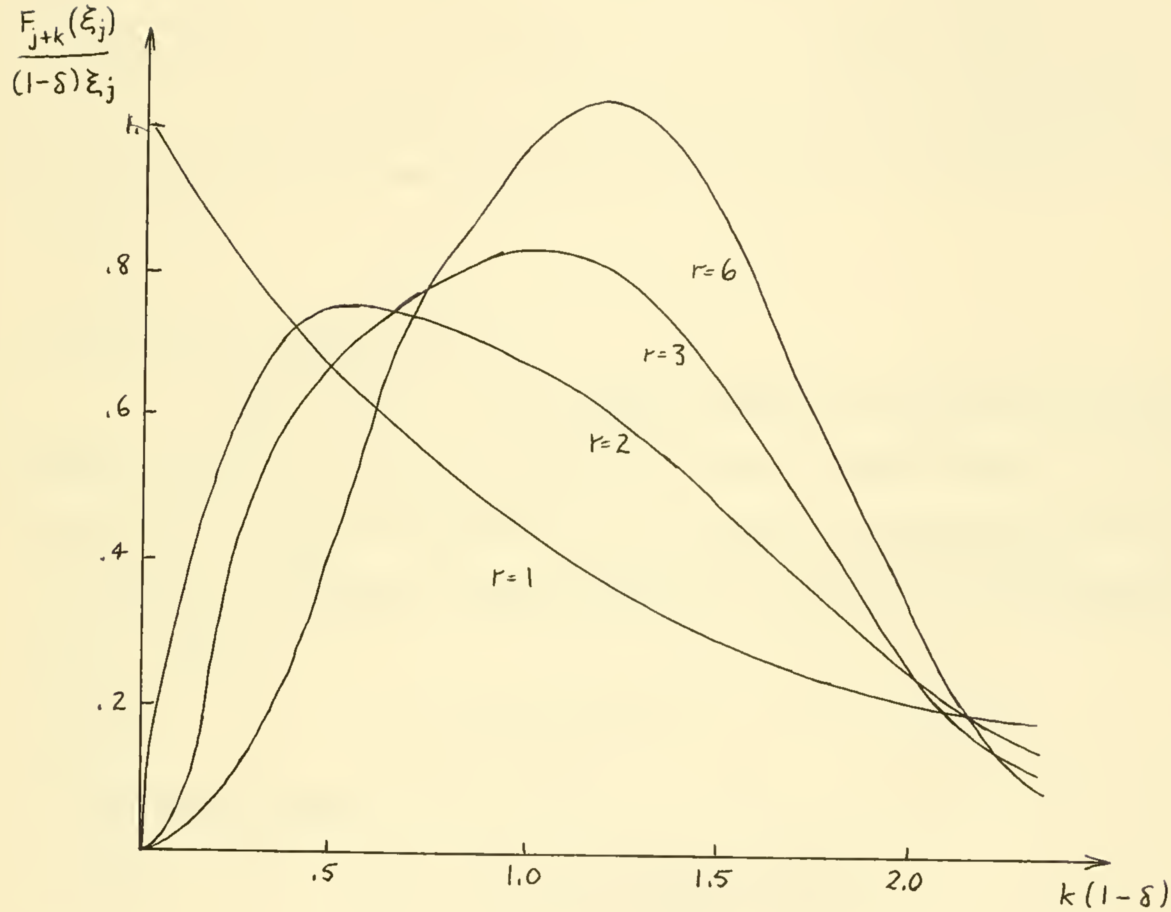
[1] Chart 5 is adapted from [24, Figure 9-7]

[2] Strictly speaking, this is of course not true for the Pascal distribution, for which $F_j(\xi_j)/(1-\delta)\xi_j$ is 1 for r = 1, $(1-\delta)$ for r = 2, $(1-\delta)^2$ for r = 3 and so forth. The continuous presentation of chart 5 provides only an approximate representation of the Pascal distribution. The remarks supra on the form of distributions defined by different values of r are consequently only qualitatively correct.

CHART 5

THE TIME PATTERN OF EFFECTS OF RANDOM SHOCKS

DEFINED BY VARIOUS VALUES OF r



random shocks.

Thus, since $k = t - j$,

(5.1.11) $$X_t = \sum_{j=0}^{t} \binom{r+t-j-1}{t-j} (1-\delta)^r \delta^{t-j} \xi_j .$$

The second stage model is thus defined in the two parameters δ and r.

In order to bring out the implications of this model, I shall first analyze γ_{FR} in the case where $r = 1$ and then turn to cases where $r > 1$. Since the marginal effects of specification of the first-stage model as in sections 4.2 and 4.3 should not differ in their qualitative aspects when applied to either second-stage model, I shall not relax the assumption that the ξ_j are a pure random process.

5.2 Behavior of γ_{FR} when $r = 1$

So long as ξ_j is a pure random process, the behavior of γ_{FR} can be easily determined from the fact that X_j is then a Markov process. As is well-known (as for instance Bartlett [4, pp. 144-145]), the autocorrelation function associated with a Markov process defined as in (5.1.7) is a decreasing exponential,[1] so that

(5.2.1) $$R_L = (R_1)^L = \delta^L .$$

Consequently, substituting in (2.5),

(5.2.2) $$\gamma_{FR} = \frac{1 - [\delta^F + \delta^R - \delta^{F+R}]}{2[1 - \delta^R]} .$$

Hence γ_{FR} is negative for $0 \leq \delta < 1$, as would be expected from the fact that the Markov process defined in (5.1.7) is stable. As would moreover be expected from the identity of the two special cases defined by $\delta = 0$ and $\lambda = 0$, γ_{FR} is equal

---

[1] Bartlett proves this for the case where $\xi_j = b\,\zeta_j$; it is however an easy extension to show that this result is also valid for $\xi_j = a + b\,\zeta_j$, $a \neq 0$.

to - 1/2 when $\delta = 0$. Perhaps more surprising is the fact that γ_{FR} converges to - 1/2 for any $\delta < 1$ as F and R are increased.

The relationship between γ_{FR} and R for various values of δ when F = R is presented in Chart 6. This chart affords some interesting comparisons with the form of $\gamma(F,R)$ defined by the analogous model of Section 3.1. For the case where a = 0, the differences in $\gamma(F,R)$ are brought out by comparing Chart 6 with Chart 1. As is clear from such a comparison, the differences are substantial. For the model of section 3, γ_{FR} goes to zero as R increases for $\lambda > 0$, while for the Markov process γ_{FR} goes to - 1/2 for $\delta < 1$ as F and R increase. This difference should be evident from the fact that, from (5.1.3),

$$\lim_{k \to \infty} F_{j+k}(\xi_j) = 0 \ . \tag{5.2.3}$$

The Markov process specification increasingly approaches that of the section 3 model for $\lambda = 0$ as F and R are increased, as a result of the increased evanescence of random shocks specified by (5.2.3).

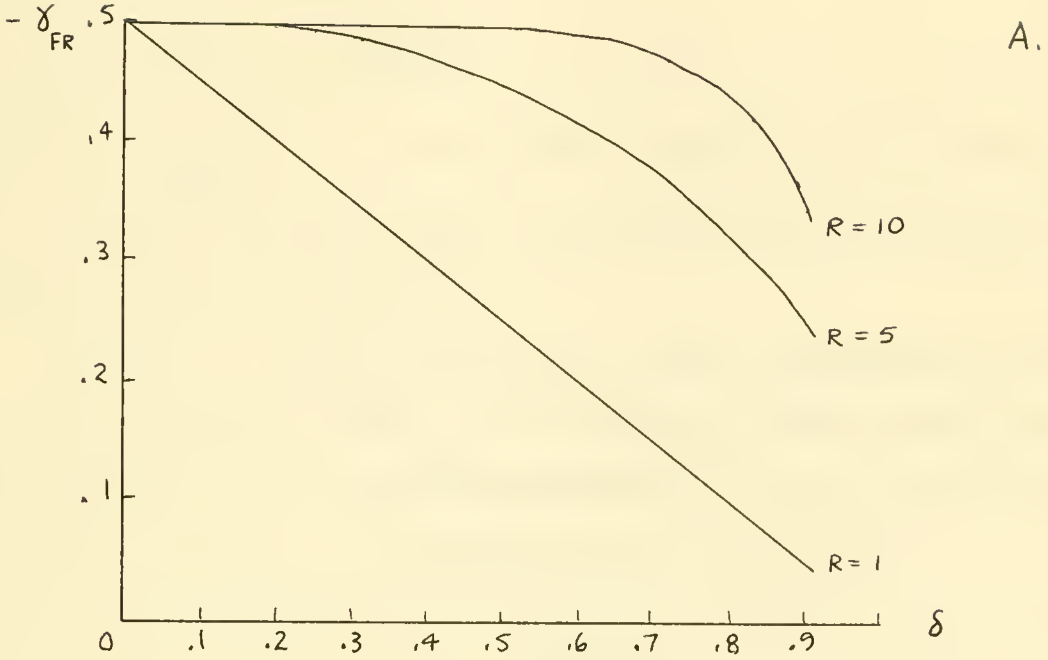
For non-zero a, the difference between the two models is even more marked. As shown in Section 4.1, the degree of regressivity in a series defined by (4.1.2) for a given value of λ decreases as $|a|$ increases. Indeed, for some combinations of λ, F, R, a and b, γ_{FR} is positive. By contrast, γ_{FR} is unchanged in the Markov process model as a departs from zero. This will be shown by computing γ_{FR} directly in the usual manner, rather than from the autocorrelation function. From (5.1.11), it is evident that

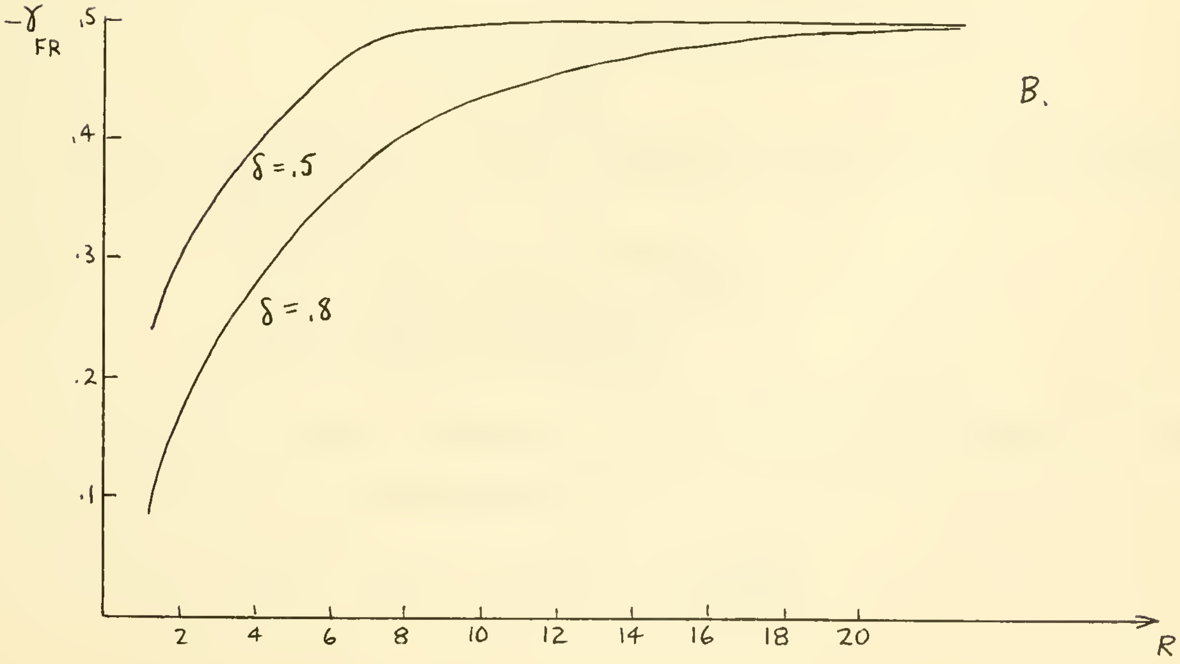
$$X_t = a(1 - \delta^{t+1}) + b\,(1 - \delta) \sum_{j=0}^{t} \delta^{t-j} \zeta_j \tag{5.2.4}$$

so that, obtaining Δ_F and Δ_R from (5.2.4), substituting in (2.3), taking expected values of product moments, and closing sums,

CHART 6

RELATIONSHIP BETWEEN γ_{FR}, R, AND δ





$$(5.2.5)\quad \gamma_{FR} = \frac{a^2H\delta^{2t+2} + b^2H\left(\frac{1-\delta}{1+\delta}\right)\left[\frac{\delta^R}{\delta^R-1}(1-\delta^{2R}) + \delta^{2R} - \delta^{2t+2}\right]}{a^2\left(\frac{\delta^R-1}{R}\right)^2\delta^{2t+2} + b^2\left(\frac{1-\delta}{1+\delta}\right)\left[1-\delta^{2R} + \frac{\delta^R-1}{\delta^R}\delta^{2R} - \delta^{2t+2}\right]}$$

where $H = (\delta^F - 1)\ \dfrac{\delta^R - 1}{\delta^R}$

Since $0 \leq \delta < 1$, $\operatorname*{plim}_{t\in T}\delta^t = 0$ so that (5.2.5) reduces to (5.2.2). Though this result may seem surprising, it follows from the strict-sense stationarity of a Markov process defined such that $\delta < 1$.

If there is a constant linear drift in E ξ, as discussed in section 4.2.2., then the resulting process is no longer strict-sense stationary and is Markov only around the trend consequently defined in X. Since from this definition $\xi_j = a + cj + b\,\zeta_j$, then, substituting for ξ_j (5.1.11),

$$(5.2.6)\qquad X_t = a(1+\delta^{t+1}) + ct\left(1 - \frac{\delta(1-\delta^t)}{t(1-\delta)}\right) + b\,(1-\delta)\sum_{j=0}^{t}\delta^{t-1}\zeta;$$

which, as before, reduces to

$$(5.2.7)\qquad X_t = a + ct + b\,(1-\delta)\sum_{j=0}^{t}\delta^{t-j}\zeta;$$

for large enough values of t. Computing γ_{FR} from (5.2.7) in the same manner as before, it is evident that

$$(5.2.8)\qquad \gamma_{FR} = \frac{c^2FR - b^2(\delta^F-1)(\delta^R-1)\,\frac{1-\delta}{1+\delta}}{c^2R^2 + 2b^2(1-\delta^R)\,\frac{1-\delta}{1+\delta}}$$

If c = 0, (5.2.8) of course reduces to (5.2.2). For non-zero values of c, however, it is apparent that γ_{FR} may be positive if

$$(5.2.9)\qquad c^2FR > b^2(\delta^F-1)(\delta^R-1)\,\frac{1-\delta}{1+\delta},$$

as would be expected from the fact that the unconditional expected values of future and past change are non-zero. In particular $E\Delta_F = cF$ and $E\Delta R = cR$. Equation (5.2.8) thus represents the ratio of $E\Delta_F$ to $E\Delta_R$ modified to allow for the variance in Δ_F and Δ_R around their expected values.

5.3 Behavior of γ_{FR} when $r > 1$

The nature of $\gamma(F,R)$ for the model defined by any positive value of r can perhaps best be analyzed by comparing it with the nature of $\gamma(F,R)$ for the special case defined by $r = 1$ and discussed in the previous section. Defining $\xi_j = a + b\zeta_j + cj$, as in the previous section, (5.1.11) can be rewritten as

(5.3.1) $$X_t = (1-\delta)^r \sum_{j=0}^{t} D(t-j)\,\delta^{t-j}\,[a + b\zeta_j + cj]$$

$$\text{where } D(k) = \binom{r+k-1}{k}$$

If $r = 1$, $D(k) = 1$ for all values of k and (5.3.1) reduces to (5.2.4).

In general,

(5.3.2) $$D(k) = \frac{(k+r-1)!}{k!\,(r-1)!}$$

For any value of r, $D(k) = 1$ if $k = 0$. If $k > 0$, $D(k) > 1$ for all $r > 1$, and, in particular,

(5.3.3) $$D(k+1) = \left(\frac{k+r}{k+1}\right) D(k) \quad .$$

It is apparent from (5.3.3) that both $D(k+1)/D(k)$ and the ratio of $D(k+1)/D(k)$ to $D(k)/D(k-1)$ approach unity as k increases for $r > 1$ and that both ratios approach unity for a given k as r approaches unity from above. Thus, for any given (positive) value of k, the extent to which $D(k)$ exceeds unity is increased as R is increased. Likewise, for any two given non-equal values of R, the difference between $D(k)$ for the higher R and $D(k)$ for the lower r increases as k is increased.

All these facts make it easy to pinpoint the difference between models resulting from differently specified values of r. Obtaining Δ_F and Δ_R from (5.3.1) and obtaining γ_{FR} in the usual manner, it turns out that

$$(5.3.4) \quad \gamma_{FR} = \frac{a^2\ EG + ac(EK + GJ) + c^2\ JK + b^2\ (M + N)}{a^2\ G^2 + 2acGK + c^2K^2 + b^2\ Q}$$

$$\text{where } E = \sum_{j=0}^{t} [D(t+F-j)\delta^F - D(t-j)]\delta^{t-j} + \sum_{j=t+1}^{t+F} D(t+F-j)\delta^{t+f-j}$$

$$G = \sum_{j=0}^{t-R} [D(t-j) - \frac{D(t-R-j)}{\delta^R}]\delta^{t-j} + \sum_{j=t-R+1}^{t} D(t-j)\delta^{t-j}$$

$$J = \sum_{j=0}^{t} j[D(t+f-j)\delta^F - D(t-j)]\delta^{t-j} + \sum_{j=t+1}^{t+F} j\ D(t+F-j)\delta^{t+F-j}$$

$$K = \sum_{j=0}^{t-R} j\ [D(t-j) - \frac{D(t-R-j)}{\delta^R}]\delta^{t-j} + \sum_{j=t-R+1}^{t} j\ D(t-j)\delta^{t-j}$$

$$M = \sum_{j=0}^{t-R} [D(t+F-j)\delta^F - D(t-j)] \left[D(t-j)] - \frac{D(t-R-j)}{\delta^R}\right]\delta^{2(t-j)}$$

$$N = \sum_{j=t-R+1}^{t} [D(t+f-j)\delta^F - D(t-j)]\ D(t-j)\delta^{2\ (t-j)}$$

$$Q = \sum_{j=0}^{t-R} \left[[D(t-j) - \frac{D(t-R-j)}{\delta^R}\right]^2 \delta^{2(t-j)} + \sum_{j=t-R+1}^{t} [D(t-j)]^2\ \delta^{2(t-j)}$$

If $r = 1$, then $D(k) = 1$ for all k and (5.3.4) reduces to (5.2.5). To evaluate (5.3.4) for $r > 1$, we can rewrite E and G as

$$(5.3.5) \quad E = D(t+1)\delta^{t+1} \sum_{k=0}^{F-1} \left[\frac{D(t+1+k)}{D(t+1)}\right]\delta^k \cong D(t+1) \left[\frac{1 - \delta^F}{1 - \delta}\right]$$

$$(5.3.6)\quad G = D(t-R+1)\delta^{t-R+1} \sum_{k=0}^{R-1} \left[\frac{D(t-R+1+k)}{D(t-R+1)}\right]\delta^{k} \cong D(t-R+1)\delta^{t-R+1} \quad \frac{1-\delta^{F}}{1-\delta}$$

where the approximations follow from the asymptotic behavior of D(K+1)/D(k) described above. It is clear from this same asymptotic behavior that the product EG will be arbitrarily close to zero if $[D(t-R+1)\delta^{t-R+1}]$ is arbitrarily close to zero. It is clear furthermore that

$$(5.3.7)\qquad \lim_{k \to \infty} \frac{D(k+1)\delta^{k+1}}{D(k)\delta^{k}} = \delta \quad .$$

Consequently

$$(5.3.8)\qquad \lim_{k \to \infty} E = \lim_{k \to \infty} G = 0$$

The rapidity of convergence of this limit is of course dependent upon δ and r. For $\delta < 0.5$, $D(k)\delta^{k}$ is less than 0.001 for $k > 20$ and $r < 7$. For most time series to be analyzed, it should be safe to assume that both E and G are arbitrarily close to zero. We thus obtain

$$(5.3.9)\qquad \gamma_{FR} = \frac{c^2 JK + b^2 (M+N)}{c^2 k^2 + b^2 Q},$$

an expression which is dependent only on c, b, k, δ, F, and R. If, moreover, $c = 0$ then $\gamma_{FR} = (M+N)/Q$ and is thus independent of the parameter of the first-stage process generating the random shocks.

6. A note on applications to objective data

In this section I shall present an illustrative application of the analysis of sections 3 through 5. In order to apply this analysis to objective data, it is of course, necessary to say something about the likelihood of obtaining estimates of γ_{FR} from a realization of a particular process which are within some stated range around the theoretical values of γ_{FR} for that process. I shall therefore briefly "say something" about this problem in section 6.1 before turning to the application in section 6.2

6.1 The evaluation of realization estimates

Since I have defined γ_{FR} as the expected value of the estimate γ_{FR} defined in equation (2.3), and since moreover γ_{FR} asymptotically converges to its expected value, it is clear that estimates of γ_{FR} obtained from realizations of a given process discussed in section 3, 4, or 5 will converge asymptotically to the values derived for that process. The likelihood of being parent to a given realization is thus higher for one process than for another if the form of γ (F,R) defined by the first process more closely accords with the estimates of γ_{FR} computed from the realization that does that defined by the second. However, the difference between the two likelihoods may be small, even if γ (F,R) is quite different for the two processes compared, if the sampling distribution of estimates for γ_{FR} is wide for realizations of the two processes containing the same number of observations as the given realizations. The fact of asymptotic convergence consequently does not of itself have much practical value, particularly since the size of realizations which can be obtained from relevant historical records is usually small. It is the small-sample distributions of estimates of γ_{FR} with which we must be concerned. Small-sample distributions of estimates

of γ_{FR} obtained by repeated simulation will be presented in [11] for various processes analyzed in sections 3, 4, and 5.

6.2 An illustrative application

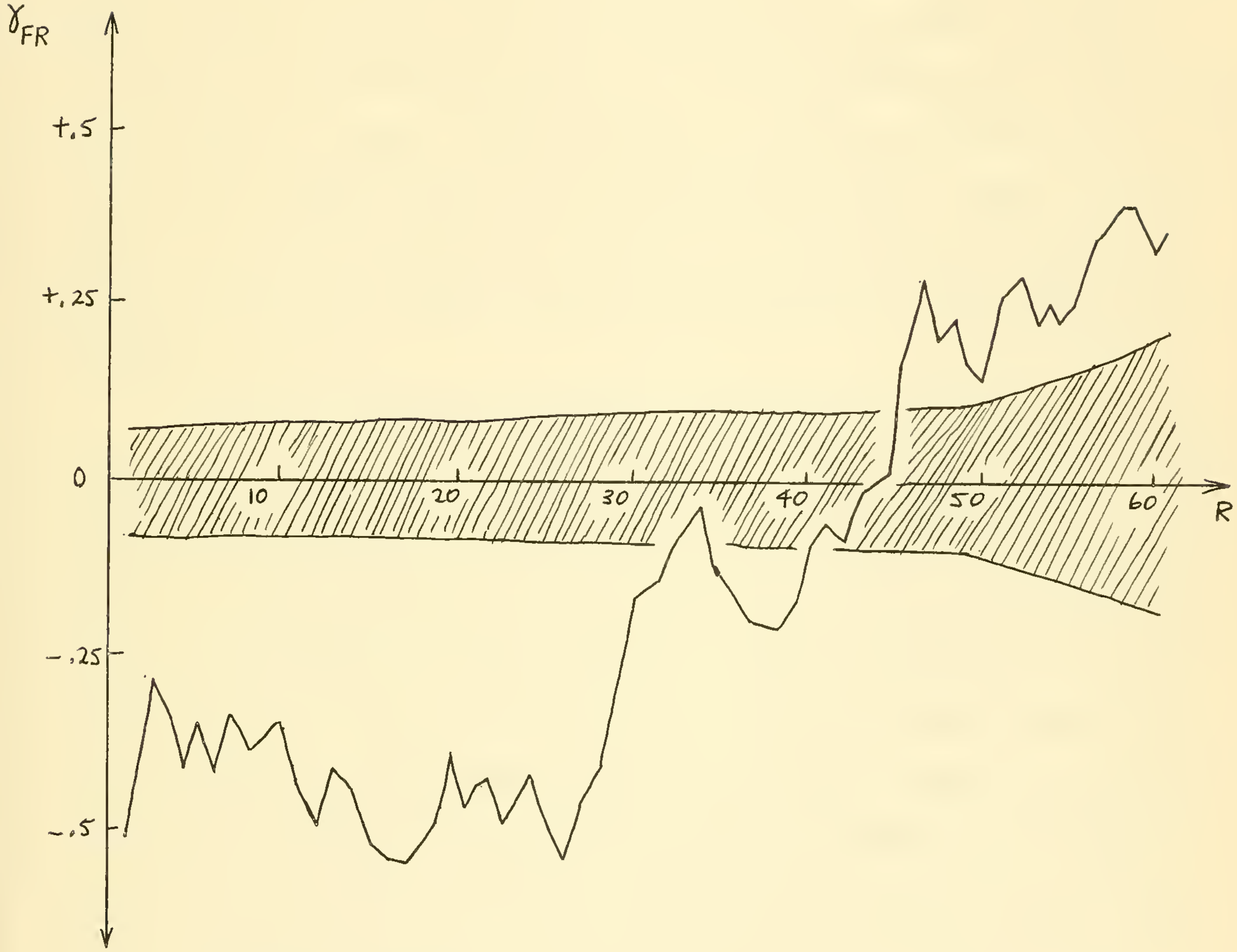
The application to be discussed in this section involves the analysis of estimates of γ_{FR} obtained from a sample of time series data rather than from subjective information. The sample data consist of a set of monthly observations of total domestic shipments of carbon black for the 14 years 1947 through 1960. Since these data are discussed in [6], I shall not further comment here, but shall instead immediately turn to an analysis of the estimates presented in [8] of γ_{FR} for cases where F = R.

The estimates of γ_{FR} obtained from the carbon black data are presented in Chart 7 for values of R ranging from 1 to 60. (F is set equal to R.) The least-squares estimates of the standard errors of estimate of each γ_{FR} are also presented in Chart 7, plotted symetrically around null values of γ_{FR}.

It is evident that a systematic pattern exists in the estimates of γ_{FR} presented in Chart 7, the most obvious characteristic of which is the positive sign of γ_{FR} for $R > 43$. Such a pattern corresponds with what a priori information would suggest: namely, that "rational" forecasts of ΔF should imply negative values of γ_{FR} in the short-term for most industrial series, reflecting the relatively large short-term variability of such data, but should imply positive values of γ_{FR} over periods long enough for secular trends to be manifested. Both a priori information and the pattern exhibited in Chart 7 make highly unlikely the possibility that the carbon black shipments are a realization of the purely random process defined by (3.1.2) with $\lambda = 0$ or by (5.1.11) with $\delta = 0$. Likewise unlikely is the possibility that carbon black shipments are a random walk, even around some trend since the asymptotic values of γ_{FR} would in that event be negative for no

CHART 7

RELATIONSHIP BETWEEN SUCCESSIVE CHANGES
IN MONTHLY SHIPMENTS OF CARBON BLACK



SOURCE: Table 2-2 in [8]

values of R.

If it could be assumed that random shocks affecting carbon black shipments are identically and independently distributed, then it would be possible to restrict analysis to the model defined by (4.1.3) because of the non-positiveness of γ (F,R) for all F and R for the model defined by (5.2.4). Actually, it is quite likely that there has been a drift in the expected value of these random shocks, making it useful to compare the extent to which the estimates of γ_{FR} for carbon black shipments accord with more complicated models. However, since the method of analysis is the same, I shall restrict myself here to comparing the estimates presented in Chart 7 with the $\gamma(F,R)$ defined by various values of the parameters of (4.1.3).

Since the estimates of γ_{FR} are all positive and more than one standard error unit away from zero for $R > 45$ and all negative and again more than one standard error unit away from zero for $R < 35$, the coefficients of variation of the random shocks which must be associated with given values of λ in order that γ_{FR} be zero for values of R within that range can be calculated from (4.1.10) again setting F = R. The values of b/a thus obtained can then be substituted into (4.1.9) along with the given values of λ to generate values of γ_{FR} to compare with every estimate of γ_{FR} obtained from the realization data. Chart 8 presents the results of a set of such comparisons for generated values of γ_{FR} using the midpoint and extreme values of the range of R for which the carbon black γ_{FR} estimates are close to zero. The standard of comparison is the overall mean squared difference between the carbon black γ_{FR} estimates and the generated values of γ_{FR}. Some of the sets of γ_{FR} values generated for selected values of λ are presented in Chart 9. As Charts 8 and 9 indicate, the "best fit" values of λ would seem to be equal to or below 0.01, which would indicate a greater degree of randomness in the shipments' series than perhaps might be expected *a priori*.

CHART 8

MEAN SQUARED DIFFERENCE BETWEEN THE ESTIMATES OF γ_{FR} PRESENTED IN CHART 7 AND VALUES OF γ_{FR} COMPUTED FROM (4.1.9) FOR F=R=35,40 and 45 AND FOR SELECTED VALUES OF λ

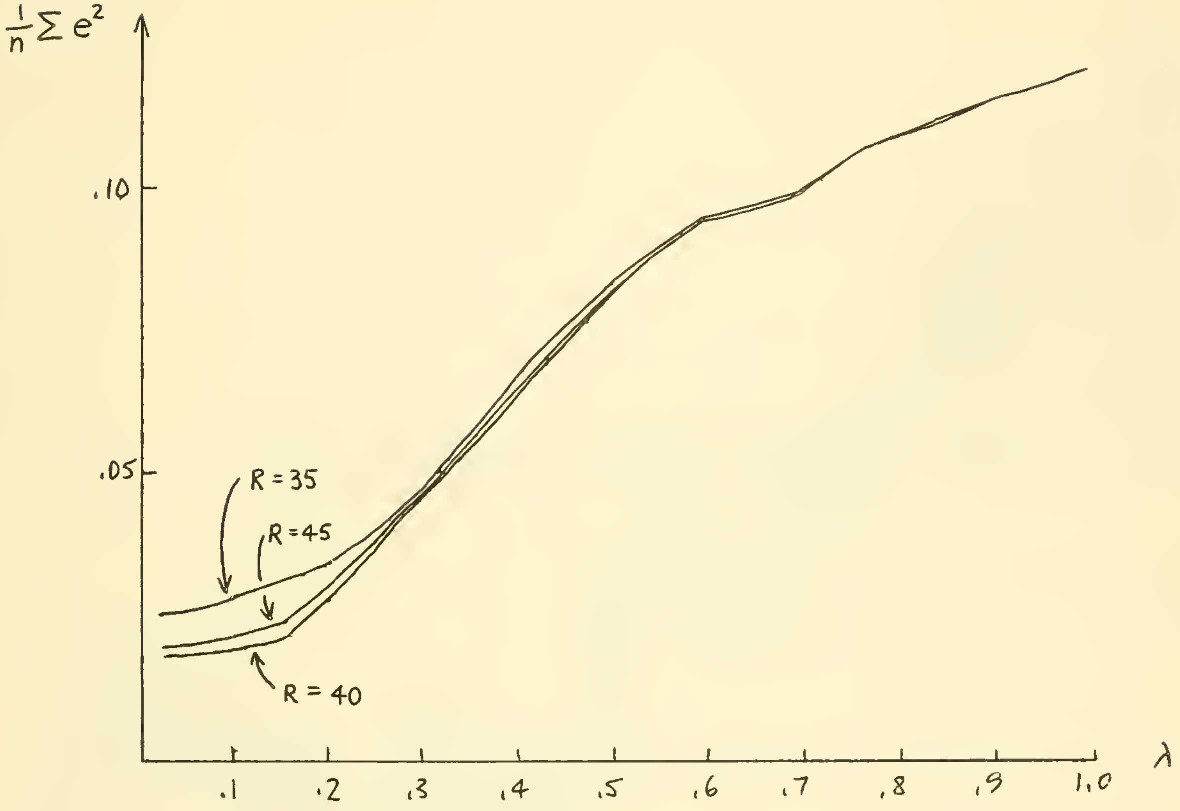
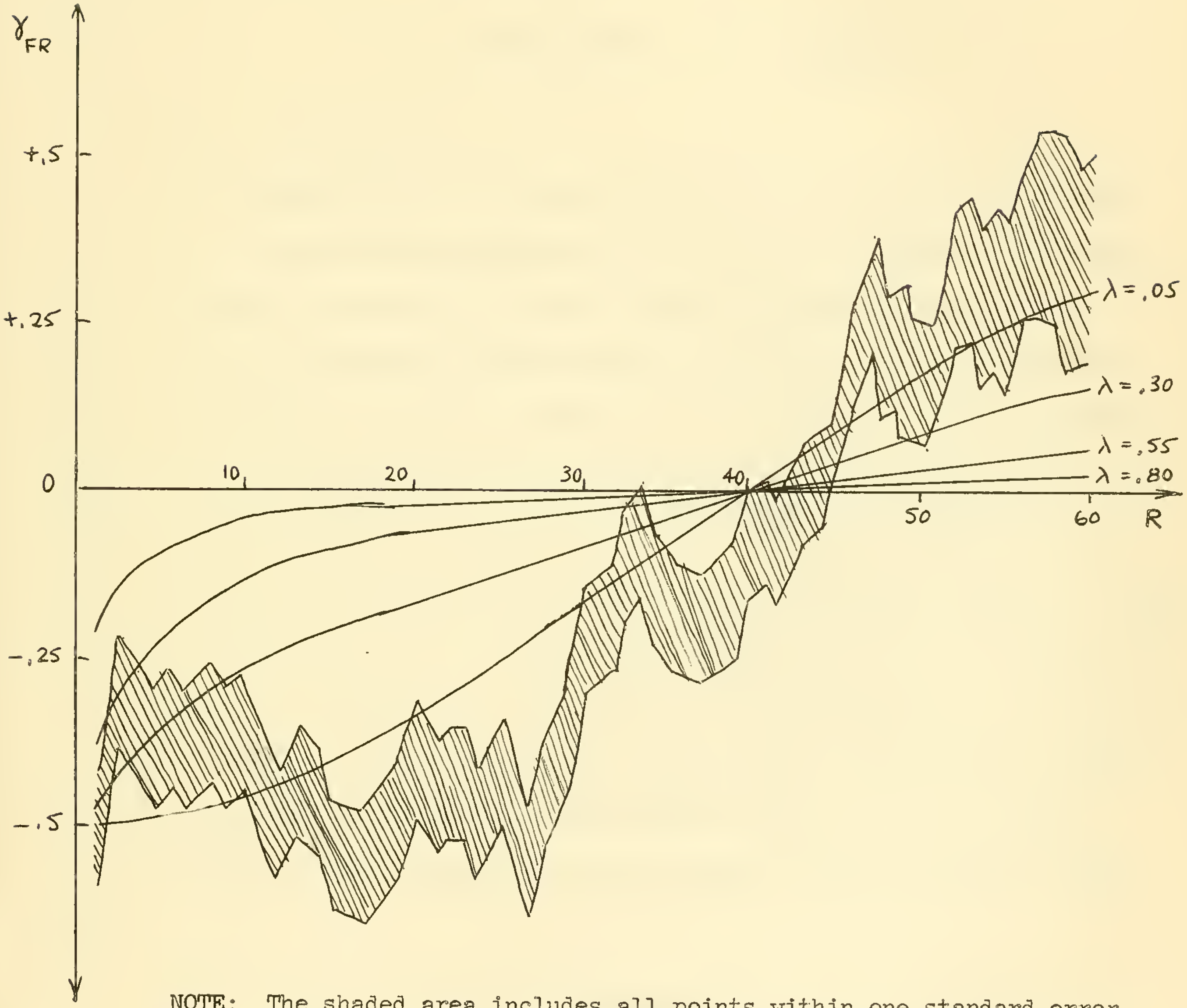


CHART 9

COMPARISON OF γ_{FR} COMPUTED FROM (4.1.9)
FOR F=R=40 AND SELECTED VALUES OF λ WITH
THE ESTIMATES OF γ_{FR} FOR CARBON BLACK
PRESENTED IN CHART 7



NOTE: The shaded area includes all points within one standard error unit (as estimated in [8]) of the estimates of γ_{FR} presented in Chart 7

However, it should be recalled that in this model $E\Delta_F = a\lambda F$. Fitting a constrained regression on F to the mean values of Δ_F for the carbon black data yields an estimate $a\lambda = 260$, so that for $\lambda = 0.01$, for example, we would obtain, substituting a, λ and V in (4.1.3),

$$X_t = 26000(1+.01t) + 104 \sum_{j<t} \zeta_j, \tag{6.2.1}$$

the variable t being undefined.

Other models would be compared with this one by using the same method of analysis and discriminating among them by mean squared-difference criteria. Further analysis of the differences could yield additional information that might be useful in specifying the time series structure of the observed data. For the model fitted in the above paragraph, for instance, analysis of the mean squared differences between observed and generated γ_{FR} over smaller ranges of R indicates that for $1 \leq R \leq 10$ this mean squared difference is lowest for the model specified by $\lambda = .30$, thus indicating a substantially lesser degree of randomness over the short run and in turn suggesting presence of some first-order autocorrelation in the random shocks ξ_j. Analysis of such models may also be extended by using estimated equations such as (6.2.1) as a means of estimating the random shocks ζ_j, and then analyzing the interdependence between the ζ_j and other factors. But this sort of analysis begins to take us beyond the scope of this paper.

7. Conclusion

In a recent monograph [23, p. 59], Frank Fisher has included an apt quotation from Heraclitus: "The Lord whose oracle is at Delphi neither reveals nor conceals, but he indicates his meaning through hints." Faced with the

problem of identifying the time-structure of economic interdependencies, the econometrician cannot but be impressed by the difficulties inherent in interpreting the hints contained in observed data. While the analysis described in this paper has attempted to unravel some of the simpler aspects of this specification problem, it has raised many more unanswered questions. But this perhaps is to be expected in the Delphic art of econometrics.

REFERENCES

1. Alexander, Sidney S. "Price Movements in Speculative Markets: Trends or Random Walks," Industrial Management Review, 2 (May 1961), pp. 7-26.

2. Arrow, Kenneth J., and Marc Nerlove. "A Note on Expectations and Stability," Econometrica, 26 (April 1958), pp. 297-305.

3. Bachelier, M. L. Theorie de la Speculation (Paris: Gauthier-Villars, 1900).

4. Bartlett, M. S. An Introduction to Stochastic Processes (Cambridge: Cambridge University Press, 1955).

5. Bartlett, M. S. "The Impact of Stochastic Process Theory on Statistics," Probability and Statistics: The Harald Cramer Volume (New York: Wiley, 1959), pp. 39-49.

6. Billingsley, Patrick. Statistical Inference for Markov Processes (Chicago: University of Chicago Press, 1961).

7. Billingsley, Patrick. "Statistical Methods in Markov Chains," Annals of Mathematical Statistics, 32 (March 1961), pp. 12-40.

8. Bossons, John. "The Nature of 'Rational' Expectations for a Manufacturer," dittoed, January 1962.

9. Bossons, John, and Franco Modigliani. "The Source of Regressiveness in Surveys of Businessmen's Short-Run Expectations," The Quality and Economic Significance of Anticipations Data (Princeton: Princeton University Press for the National Bureau of Economic Research, 1960), pp. 239-259.

10. Bossons, John, and Franco Modigliani. "On the Reasonableness of Regressive Expectations," Part 1, dittoed, June 1961.

11. Bossons, John, and William Steiger. "The Sampling Properties of Parameters of Least Squares Regressions between Contiguous Changes in the Time Series: A Simulation Analysis," in process.

12. Brown, Robert G. Statistical Forecasting for Inventory Control (New York: McGraw-Hill, 1959).

13. Brown, Robert G., and Richard F. Meyer. "The Fundamental Theorem of Exponential Smoothing," Operations Research, 9 (October 1961), pp. 673-685.

14. Cagan, Phillip. "The Monetary Dynamics of Hyperinflation," in Studies in the Quantity Theory of Money (Chicago: University of Chicago Press, 1956), pp. 25-117.

15. Cootner, Paul. "Random versus Systematic Patterns in Stock Price Movements," paper delivered before the 1961 Annual Meeting of the American Finance Association 28 December 1961.

16. Cramer, Harald. Mathematical Methods of Statistics (Princeton: Princeton University Press, 1946).

17. Davis, Michael C. "Optimum Systems in Multi-Dimensional Random Processes," Sc. D. Thesis (Cambridge: Massachusetts Institute of Technology, 1961).

18. Davis, Michael C. "Wiener Theory Revisited in N Dimensions," Multilithed, 1961.

19. Durbin, J. "The Fitting of Time-Series Models," Revue de l-Institut International de Statistique, 28 (1960), pp. 233-243.

20. Eisner, Robert. "A Distributed Lag Investment Function," Econometrica, 28 (January 1960) pp. 1-29.

21. Ellsberg, Daniel. "Risk, Ambiguity, and the Savage Axioms," Quarterly Journal of Economics, 75 (November 1961), pp. 643-669.

22. Feller, William. An Introduction to Probability Theory and its Applications, Volume 1, revised edition (New York: Wiley, 1957).

23. Fisher, Franklin M. A Priori Information and Time Series Analysis: Essays in Economic Theory and Measurement (Amsterdam: North-Holland, 1962).

24. Forrester, Jay W. Industrial Dynamics (Cambridge: The Massachusetts Institute of Technology Press, 1961).

25. Friedman, Milton. "The Methodology of Positive Economics," Essays in Positive Economics (Chicago: University of Chicago Press, 1953), pp. 3-43.

26. Friedman, Milton. A Theory of the Consumption Function (Princeton: Princeton University Press for the National Bureau of Economic Research, 1957).

27. Friedman, Milton, and Simon Kuznets. Income from Independent Professional Practice (New York: National Bureau of Economic Research, 1945).

28. Frisch, Ragnar. "Propagation Problems and Impulse Problems in Dynamic Economics," Economic Essays in Honor of Gustav Cassel (London, 1933), pp. 171-205.

29. Grenander, Ulf, and Murray Rosenblatt. Statistical Analysis of Stationary Time Series (New York: Wiley, 1957).

30. Hald, A. The Decomposition of a Series of Observations Composed of a Trend, a Periodic Movement, and a Stochastic Variable (Copenhagen: Gads Forlag, 1948).

31. Harary, Frank, and Benjamin Lipstein. "The Dynamics of Brand Loyalty: A Markovian Approach," Operations Research, 10 (January 1962), pp. 19-40.

32. Herniter, Jerome D., and John F. Magee. "Customer Behavior as a Markov Process," Operations Research, 9 (January 1961), pp. 105-122.

33. Holt, Charles C. "Forecasting Trends and Seasonals by Exponentially Weighted Moving Averages," O. N. R. Memorandum No. 52 (Carnegie Institute of Technology, April 1957) dittoed.

34. Holt, Charles C., Franco Modigliani, and John F. Muth, and Herbert A. Simon. Planning Production, Inventories, and Work Force (Englewood: Prentice Hall, 1960).

34a. Houthakker, Hendrik S. "Systematic and Random Elements in Short-Term Price Movements," American Economic Review, 51:2 (May 1961), pp. 164-172.

35. Kendall, Maurice G. "The Analysis of Economic Time Series--Part I: Prices." Journal of the Royal Statistical Society (Series A), 96 (1953), pp. 11-25.

36. Kerrich, J. E. "Random Remarks," The American Statistician, 15 (June 1961), pp. 16-20.

37. Koyck, Leendert M. Distributed Lags and Investment Analysis (Amsterdam: North-Holland, 1954).

38. Kuh, Edwin. "Capital Stock Growth--A Microeconometric Study," dittoed, September 1961.

39. Lintner, John. "Distribution of Incomes of Corporations among Dividends, Retained Earnings, and Taxes," American Economic Review: Papers and Proceedings, 46 (May 1956), pp. 97-113.

40. Maffei, Richard B. "Brand Preferences and Simple Markov Processes," Operations Research, 8 (March 1960), pp. 210-218.

41. Magee, John F. Production Planning and Inventory Control (New York: McGraw-Hill, 1958).

42. Mandelbrot, Benoit. "The Pareto-Levy Law and the Distribution of Income," International Economic Review, 1 (May 1960), pp. 79-106.

43. Mandelbrot, Benoit. "Stable Paretian Random Functions and the Multiplicative Variation of Income," Econometrica, 29 (October 1961), pp. 517-543.

44. Mandelbrot, Benoit. "Paretian 'Brownian Motion' Models of The Fluctuations of Commodity Prices," IBM Research Report No. RC-470.

45. Maverick, Lewis. Time Series Analysis: Smoothing by Stages (San Antonio: Anderson, 1945).

46. Mills, Edwin S. "The Use of Adaptive Expectations in Stability Analysis: A Comment," Quarterly Journal of Economics, 75 (May 1961) pp. 330-335.

47. Muth, John F. "Optimal Properties of Exponentially Weighted Forecasts of Time Series with Permanent and Transitory Components," Journal of the American Statistical Association, 55 (June 1960), pp. 299-306.

48. Muth, John F. "Rational Expectations and the Theory of Price Movements," Econometrica, 29 (July 1961), pp. 315-335.

49. Muth, John F. "Distributed Leads," paper presented at the 1961 New York meetings of the Econometric Society.

50. Nerlove, Marc. "A Note on Long-Run Automobile Demand," Journal of Marketing -- (July 1957), pp. 57-64.

51. Nerlove, Marc. Distributed Lags and Demand Analysis for Agricultural and Other Commodities, U. S. Department of Agriculture, June 1958.

52. Nerlove, Marc. The Dynamics of Supply: Estimation of Farmers' Response to Price (Baltimore: John Hopkins, 1958).

53. Nerlove, Marc. "Adaptive Expectations and Cobweb Phenomena," Quarterly Journal of Economics, 73 (May 1958), pp. 227-240.

54. Parzen, Emmanuel. Stochastic Processes with Applications to Science and Engineering (San Francisco: Holden-Day, 1961), preliminary edition.

55. Parzen, Emmanuel. "An Approach to Time Series Analysis," Annals of Mathematical Statistics, 32 (December 1961), pp. 951-989.

56. Raiffa, Howard, "Comment," Quarterly Journal of Economics, 75 (November 1961), pp. 690-694.

57. Raiffa, Howard, and Robert Schlaifer. Applied Statistical Decision Theory (Boston: Harvard Business School Division of Research, 1961).

58. Roberts, Harry. "Stock Market 'Patterns' and Financial Analysis," Journal of Finance, 14 (March 1959), pp. 1-10.

59. Rosenblatt, Murray, "Statistical Analysis of Stochastic Processes," Probability and Statistics: The Harald Cramer Volume (New York: Wiley, 1959), pp. 248-275.

60. Random Processes (New York: Oxford University Press, 1962).

61. Schlaifer, Robert. Introduction to Statistics for Business Decisions (New York: McGraw-Hill, 1961).

62. Shubik, Martin. Strategy and Market Structure: Competition, Oligopoly and the Theory of Games (New York: Wiley, 1959).

63. Slutzky, Eugen. "The Summation of Random Causes as the Source of Cyclic Processes," Econometrica, 5 (April 1937), pp. 105-146. A translation with revisions of a paper appearing in Russian in Problems of Economic Conditions (Moscow), 3:1 (1927).

64. Solow, Robert M. "On a Family of Lag Distributions," Econometrica, 28 (April 1960), pp. 393-406.

65. Stigum, Bernt. "Dynamic Stochastic Processes," typescript, September, 1961.

66. Stone, Richard, and D. A. Rowe. "Aggregate Consumption and Investment Functions for the Household Sector Considered in the Light of British Experience," National Okonomisk Tidskrift, (1956), pp. 1-32.

67. Theil, Henri. Economic Forecasts and Policy (Amsterdam: North-Holland, 1958).

68. Theil, Henri, and R. M. Stern. "A Simple Unimodel Lag Distribution," Metroeconomica, 12 (1960) pp. 111-119.

69. Whittle, Peter. "Some Recent Contributions to the Theory of Stationary Processes," Appendix 2 in Wold [72].

70. Wiener, Norbert. Nonlinear Problems in Random Theory (Cambridge: The Technology Press, 1958).

71. Winters, P. R. "Forecasting Sales by Exponentially Weighted Moving Averages," Management Science, 6 (1960), pp. 324-342.

72. Wold, Herman. A Study in the Analysis of Stationary Time Series, second edition (Stockholm: Almqvist and Wiksell, 1954).

73. Working, Holbrook. "A Random Difference Series for Use in the Analysis of Time Series," Journal of the American Statistical Association, 29 (March 1934), pp. 11-24.

74. Yule, G. Udny. "On a Method of Investigating Periodicity in a Disturbed Series," Transaction of the Royal Society, London, Series A, 226 (1927), pp. 267-298.